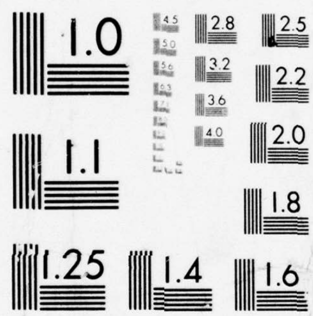
AD-A036 463 AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OHIO SCH--ETC F/G 5/3
AN INVESTIGATION OF THE FORECASTING OF AERONAUTICAL PRICE INDIC--ETC(U)
DEC 76 T P LENNERTZ

UNCLASSIFIED GOR/SM/76D-9 NL

1 OF 2
AD
A036463



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

UNITED STATES AIR FORCE

AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

GOR/SM/76D-9

AN INVESTIGATION OF THE FORECASTING OF AERONAUTICAL PRICE INDICES

THESIS

Presented to the Faculty of the School of Engineering
of the Air Force Institute of Technology

Air University

in Partial Fulfillment of the
Requirements for the Degree of

Master of Science

by

Thomas P. Lennertz, B.S.
Captain USAF

Graduate Operations Research

December 1976

(see 1473)

GOR/SM/76D-9

AN INVESTIGATION OF THE FORECASTING OF AERONAUTICAL PRICE INDICES

THESIS

GOR/SM/76D-9 Thomas P. Lennertz
Captain USAF

Approved for public release; distribution unlimited.

Preface

Dr. Keith Womer first interested me in the aircraft price index predictions annually done by the Aeronautical Systems Division (ASD) of the United States Air Force Systems Command. Later I learned that Dave Benoy of ASD was interested in new techniques that might improve the price index predictions. After talking to Dave I saw an opportunity to complete a thesis that not only would be useful to the Air Force, but also would use some of the knowledge I learned in my graduate courses. The thesis did investigate new prediction techniques and provide prediction intervals.

I wish to extend my thanks to Dr. Keith Womer, my thesis advisor, and to Major Saul Young and Dr. Leonard Gaston, my readers, and to Charlette Kjesbo, my typist. Special thanks goes also to Dave Benoy who was extremely patient and helpful in describing his work.

My deepest thanks goes to my wife, Robin, who not only helped type drafts of the thesis but also endured six months of a grumpy, nervous, and seldom seen husband.

Thomas P. Lennertz

Contents

List of Figures

List of Tables

Abstract

An investigation of the forecasts of aeronautical price indices done by the Aeronautical Systems Division (ASD) of the United States Air Force Systems Command was performed. The state of the art in correction for autocorrelation was summarized. The historical price indices used by ASD were recalculated. Using the recalculated historical composite price indices, different forecasting techniques were examined including: simple linear regression, multivariate regression, regression with nonlinear variables, Generalized Least Squares with the first order autoregressive model, first differences, and per cent change regression. The Gross National Product price deflator, the Primary Metals Index, the Electrical Machinery Index, the national total unemployment rate, and the money supply, M1, were all tried as independent variables in the regressions. The Durbin Two Step, the Search technique, and a Bayesian technique were all used to estimate the first order autoregressive coefficient, ρ. When starting with the composite price indices, the per cent change regression, with the Gross National Product price deflator as the only independent variable, was recommended as the best forecasting technique. Prediction intervals were calculated. The effect of errors in the Wharton Econometric Forecasting Associates prediction of the Gross National Product price deflator was discussed.

AN INVESTIGATION OF THE FORECASTING OF AERONAUTICAL PRICE INDICES

I. Introduction

In 1973 the Aeronautical Systems Division (ASD) of the United States Air Force Systems Command begin publishing annual reports of forecasted aeronautical price indices. Included in the reports were airframe, engine and avionics indices for both production and development. These forecasted indices were used by budget estimators in ASD to convert budget estimates in current dollars to budget estimates in then-year dollars. For example, a budget estimator in an ASD System Program Office might estimate that the budget for a specific aircraft avionics production program would be $100,000 in 1976 dollars. If one assumes that a forecasted avionics production price index for 1977 was 110 with a base of 100 in year 1976, then a budget estimator would submit a budget for avionics production of $110,000 in 1977 dollars. This budget would then be sent to higher levels for approval.

This thesis deals directly with the ASD reports of forecasted price indices. The thesis covers two major areas: one, the accuracy of the predictions, and two, the possibility of improved accuracy through different or advanced statistical techniques. The accuracy is checked by starting at the beginning with a recalculation of the historical indices, and then working all the way through to an estimate of the prediction intervals. Different statistical techniques are

examined including the addition of extra explanatory variables and the assumption of a first order autoregressive model. All mathematical techniques are clearly explained and important computer programs are listed.

II. Background

Aeronautical Systesm Division (ASD) has published annual reports of forecasted price indices since 1973. The first two reports, denoted 110A and 110B, used a similar methodology. The 110C report used a slightly different technique (Ref 1; Ref 2; Ref 3).

110A and 110B Reports

ASD used a straightforward three step methodology in their first two reports, 110A and 110B. First, they gathered historical data and developed six price indices. Second, they regressed these six price indices against the Gross National Product (GNP) price deflator. Third, they used a prediction of the GNP price deflator to forecast the price indices

In step one, ASD used historical data from aerospace contractors, combined with nationally published statistics. To develop the six composite historical indices, Production Engine, Production Airframe, Production Avionics, Development Engine, Development Airframe, and Development Avionics, ASD used a weighted average of subindices. For example, to construct the Development Engine index ASD used the following procedure. First, they constructed a purchased parts subindex using the following formula:

$$(.6)\ \text{Engine Labor Subindex} + (.4)\ \text{Raw Materials Subindex} = \text{Purchased Parts Subindex} \quad (1)$$

Second, they used this purchased parts subindex along with other subindices to form the Development Engine index.

$$(.4484)\ \text{Engine Labor Subindex} + (.1632)\ \text{Engineering Labor Subindex}$$
$$+ (.1942)\ \text{Raw Materials Subindex} + (.1942)\ \text{Purchased Parts}$$
$$= \text{Development Engine Index} \qquad (2)$$

The subindices were developed using data from the Bureau of Labor Statistics and industry. Engine labor data were from Bureau of Labor Statistics code 3722. The Raw materials index was derived from several wholesale price index categories. Engineering labor data was obtained from aerospace contractor data. The weights and methodology used were based on contractor data, and on a Rand Corporation report published in 1970 (Ref 4).

In step two, ASD related the historical data to the historical Gross National Product price deflator (GNP_d). They did this by using simple linear regression. ASD then had six regression equations, engine, airframe, and avionics for both production and development. The results for the Development Engine index were:

$$\text{Development Engine Index} = -10.8 + 1.073\ GNP_d \qquad (3)$$

R^2, the coefficient of determination, was .982. Using the criterion of R^2, it appeared that the regression had resulted in a very good fit of the data. Actually, as will be seen later, the data was autocorrelated and the fit was not nearly

as good as was first indicated by the high R^2.

In step three, ASD attempted to find a reliable predictor of the GNP deflator which could be inserted into the regression equations in order to predict the appropriate index. ASD chose the Wharton Econometric Forecasting Associates, Inc. prediction of the GNP deflator. ASD did not use Wharton's raw predictor, but smoothed the predictor before using it. ASD smoothed the predictor by fitting an exponential curve to it. ASD said they did this to reduce the impact of short term fluctuations (Ref 2:47). That then is the basic technique used by ASD in their first two reports, 110A and 110B.

110C Report

The 110C report, published in 1976, had three major changes. The first change was that ASD regressed where possible each subindex against what appeared to be an appropriate independent variable. Three independent variables were used: GNP price deflator, Primary Metals Price Index, and the Electrical Machinery Price Index. The second major change was that a per cent change regression was used in the attempt to overcome apparent autocorrelation. Per cent changes in each dependent variable were regressed against per cent changes in the independent variable. Third and last, the unsmoothed Wharton GNP deflator was used. As a result, instead of three major indices regressed for both production and development for a total of six regression equations, all but the

overhead subindices were regressed for a total of nine regression equations. These regression equations were used to predict future values of the subindices. The overhead subindices were predicted using the historic compound rate. For example, the historical 7.3% Overhead subindex was calculated by compounding 7.3% back and forth from a 1967 base value of 100. The 7.3% Overhead subindex for 1975 was

$$(100)(1 + .073)^{8} = 175.71 \qquad (4)$$

In the 110C report, this same compound rate was used to predict future values of the index. Thus the 7.3% subindex for 1976 was

$$(100)(1 + .073)^{9} = 188.54 \qquad (5)$$

Weights were applied to the subindex predictions to develop the six composite indices. The methodology of the 110A, 110B and 110C reports is contrasted in Fig. 1.

Future Interests

The per cent change regression used in report 110C apparently corrected the data for autocorrelation. ASD, however, has expressed an interest in other regression techniques that are simple, that correct for autocorrelation, and that are more effective than the per cent change regression.

In the 110C report, the subindex was first regressed, the predictions were made, and then the predictions were weighted and aggregated to form the composite indices. In

Methodology

110A and 110B	110C
1. Form historical subindices. Weight these subindices to form six composite indices.	1. Form historical subindices.
2. Regress each of the six composite indices against the historical GNP price deflator.	2. Using per cent changes, regress all of the subindices, except overhead and material, against the GNP price deflator. Regress the materials subindices against either the Primary Metals index, or the Electrical Machinery index.
3. Use the smoothed predictions of the GNP price deflator in the regression equations to predice the future values of all six indices.	3. Project the overhead subindices into the future using the historical compound rate.
	4. Use unsmoothed predictions of the independent variables in the regression equations to predict per cent changes in the subindices.
	5. Change per cent changes into subindices.
	6. Weight subindices to produce the predicted values for the six composite indices.

Fig. 1. 110A, 110B, and 110C Methodologies

the previous reports, the subindices were first aggregated and then the resulting composite indices were regressed. ASD has expressed an interest in the comparison of two techniques, both using per cent change regressions. The first technique would be to predict then aggregate, and the second technique would be to aggregate then predict.

This thesis will investigate two areas in which ASD has expressed interest:

1) Are there other techniques that can and should be used to correct for autocorrelation?

2) Which is better, regress, then aggregate, or aggregate, and then regress?

Thesis Outline

A recalculation of the historical indices was done in Chapter IV. In Chapter V the per cent change regression as done in the 110C report was repeated with one addition, prediction intervals were given. Also in Chapter V, two different approaches to per cent change regression were discussed. The first approach was to regress, and then aggregate the predictions. The second approach was to aggregate the data, and then regress. Chapter VI investigated other regression techniques. Chapter VII discussed the error in the Wharton predictors and its possible effect on the error of the predicted indices. Finally Chapter VIII summarized, presented significant findings, and made recommendations for future study.

III. Regression Theory

Preface

This chapter has been prepared in order to furnish a concise, useful, and accurate summary of applicable mathematical theory. This chapter first provides a discussion of the assumptions made when doing "ordinary" regression. It then provides a detailed discussion of autocorrelation, a violation of one of the assumptions of regression analysis.

Assumptions of the General Linear Model

The assumptions of the General Linear Model (sometimes known as Ordinary Least Squares, OLS) can be placed in two sets. The first set allows for efficient estimates. The second set, which contains only one assumption, allows for easy computation of statistical tests on the model and on its predictions. The first set of assumptions, along with brief explanations, are listed below:

1) $Y_{nx1} = X_{nxk}\beta_{kx1} + \varepsilon_{nx1}$

Y is expressed as a linear combination of k, β's with an error term, ie., the model is correctly specified.

2) $E(\varepsilon_{nx1}|X_{nxk}) = 0_{nx1}$

The expected value of the error is zero.

3) $V(\varepsilon_{nx1}|X_{nxk}) = \sigma^2 I_{nxn}$

The errors are independent of each other and all have the same variance.

4) X_{nxk} has rank k

The columns of X are independent, ie., the X variables are independent of each other.

5) X_{nxk} may consist of either fixed or random variables

where A_{bxc} stands for a matrix A, b rows by c columns (Ref 5: 111).

When these assumptions are satisfied, Ordinary Least Squares will yield the best linear unbiased estimates of the coefficients, β_{kx1} (Ref 5:119). One other assumption is frequently made.

6) Probability $(Y|X_{nxk}\beta_{kx1})$ is normally distributed. The errors are normally distributed.

This assumption means that when OLS is performed, estimates of the coefficients, β_{kx1}, will be maximum likelihood estimates. Also confidence intervals, prediction intervals, and statistical tests can be calculated.

Checks on the Assumptions of Ordinary Least Squares

In this thesis, both visual and statistical tests were performed to check on the applicability of the six assumptions or Ordinary Least Squares. OMNITAB, the computer program language used in this thesis, provided four plots that

were used to visually check assumptions 1, 3 and 6 (Ref 6: 138-146). The four plots were 1) residuals versus time, 2) residuals versus the predicted values, 3) residuals versus the independent variable, and 4) a probability plot of the residuals. These plots were utilized in the following way. If the points on plot 1 were not scattered around zero, then the errors may not have been independent and assumption 3 was assumed to be violated. If the points on plot 2 were not scattered around zero, assumption 1 was assumed to be violated, an indication that a different model such as a curvilinear regression might have done a better job. If the points on plot 3 were not scattered around zero, the variances of the errors may not have been constant, and assumption 3 was assumed to be violated. If the points on plot 4 did not lie on a straight line, then assumption 6 was assumed to be violated, an indication that the errors may not have been normally distributed.

In addition to these visual tests, numerical tests were also used to check on the assumptions. A t test was used to test the significance of the coefficients in the regression and to see if assumption 1 may have been violated (Ref 5:138). The OMNITAB program also provided estimates of the accuracy of its calculations, which were useful when assessing the independence of the columns of the X matrix, assumption 4 (Ref 6:149).

The last numerical test used the Durbin Watson statistic (Ref 5:199-201). The test checks to see if the errors

are not independent, but rather are related to each other in time, a violation of assumption 3. If the error terms are positively related to each other in time, then they are said to be positively autocorrelated. If the error terms are negatively related to each other in time, they are said to be negatively autocorrelated. This last numerical test is a statistical test consisting of a null hypothesis, H_0, an alternative hypothesis, H_A, a test statistic, D, and a rule. The test for positive autocorrelation is:

H_0 : No positive autocorrelation

H_A : Positive autocorrelation

$$D = \frac{\sum_{\alpha=1}^{n-1} (e_{\alpha+1} - e_{\alpha})}{\sum_{\alpha=1}^{n} e_{\alpha}^{2}} \qquad (6)$$

Rule	if $D < D_L$	Reject H_0
	if $D_L \leq D \leq D_U$	Inconclusive
	if $D > D_U$	Do not reject H_0

where e_α are the residuals from Ordinary Least Squares. The test for negative autocorrelation is the same with 4-D replacing D. D_L and D_U were computed and tabled by Durbin and Watson.

It is important to realize that when $D > D_U$ some authors state that the user should accept H_0 (Ref 7:358). Strictly speaking, this is not correct. There is little one can say

statistically when $D > D_U$. The significance levels given in the D_L, D_U table apply only to the rejection of H_0. In practice, however, when D approaches 2.0 the user can be fairly confident that the errors are not autocorrelated.

Assumption 2 was indirectly tested when the assumption of the correctness of the model was tested. Assumption 5 was not tested. The historical independent variables were assumed fixed. The treatment of the error in the predicted values of the independent variables is included in Chapter VII.

What to do When Autocorrelation is Detected

There are many things the analyst can do when autocorrelation is detected. He can ignore it and perform Ordinary Least Squares as usual, or he can check to see if the original model is misspecified, or he can assume a first order autoregressive model. If the user assumes a first order autoregressive model, there are many techniques available for use. All of these approaches are discussed below.

Ignore

The analyst can ignore the warning of possible autocorrelation and perform Ordinary Least Squares on the data. This will yield estimates of the coefficients; but, if the data is autocorrelated, the estimates may not be the best estimates available. Also if the data is autocorrelated, Ordinary Least Squares will yield calculated variances that underestimate the true variances (Ref 5:254-257). Fig. 2

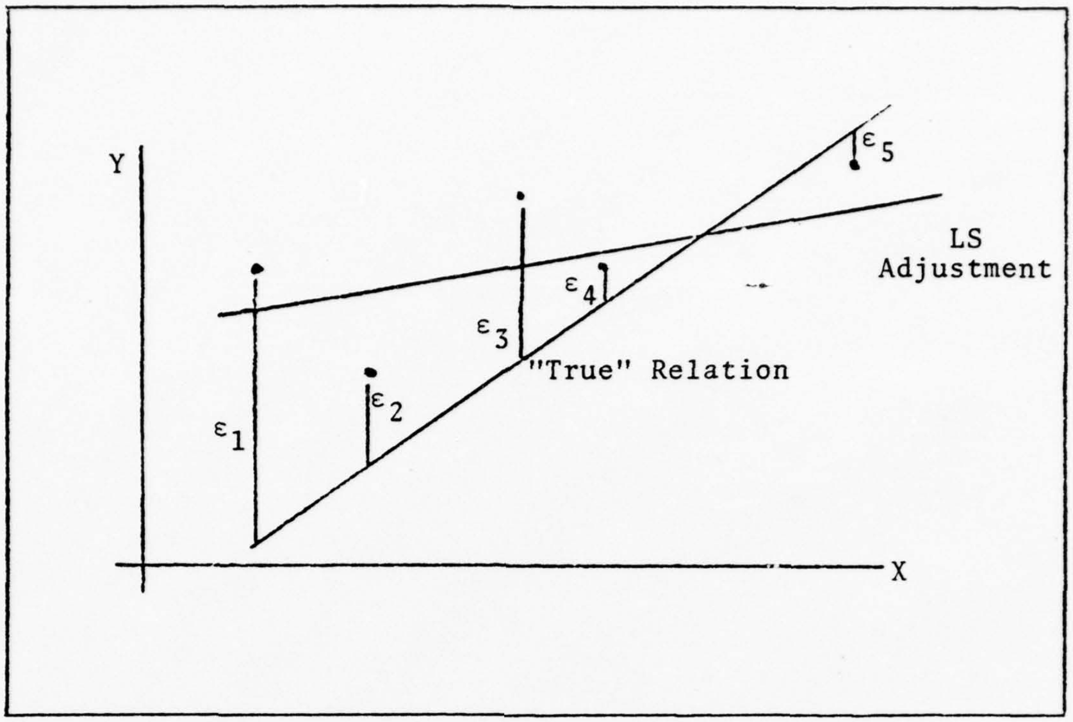


Fig. 2 Bias of the LS variance estimator under conditions of positive autocorrelation (From Ref 5:257)

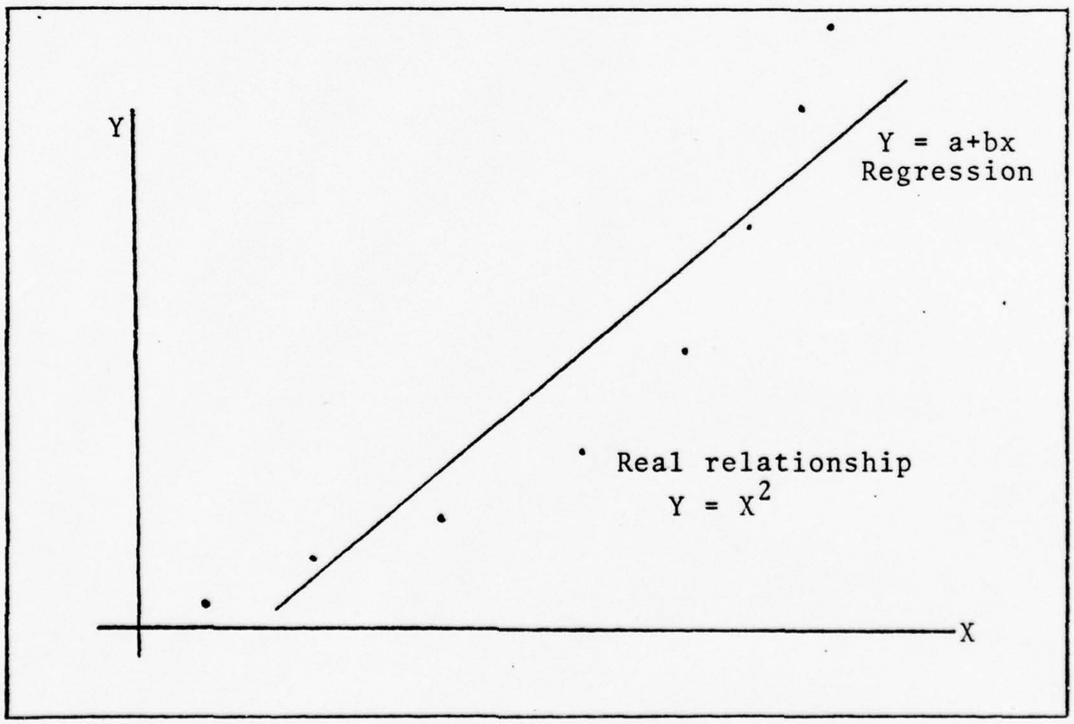


Fig. 3 Misspecified Model

is an example of what might happen if Ordinary Least Squares is performed on autocorrelated data.

Misspecified Model

A low Durbin Watson test statistic indicating autocorrelation may result from a misspecification of the model. A variable may be missing or a variable may be in the wrong form. For example if the real relationship that exists between Y and X is $Y = X^2$, and a straight line is fitted to the data, and if time is highly related to X, then the data may appear to be positively autocorrelated. The real problem is not autocorrelation, but rather a misspecification of the model. See Fig. 3.

Assume First Order Autoregressive Model

If one is not satisfied with ignoring the detection of possible autocorrelation, and if different models and additional variables have been tested to no avail, then one recommended next step is to assume a first order autoregressive model. The first order autoregressive model is:

$$Y_{nx1} = X_{nxk}\beta_{kx1} + \varepsilon_{nx1} \tag{7}$$

$$\varepsilon_t = \rho\varepsilon_{t-1} + \mu_t \tag{8}$$

where μ_t is a normally distributed independent random variable with zero mean and constant variance. ε_t is the error at time t (Ref 5:250). In other words, the error this period is equal to a constant times the error last period plus a random error. Usually another assumption is included in the

first order autoregressive model. It is that $|\rho|<1$ (Ref 5: 250). In this thesis, however, no restriction will be put on ρ. Techniques will be discussed that assume $\rho = 1$, $|\rho|<1$, and $-\infty<\rho<\infty$.

$|\rho|<1$ Assumption

The assumption of $|\rho|<1$ allows for cases of first order autocorrelation where the error does not go to infinity in time. Without going into a lot of detail, it is sufficient to say that when ρ is known a matrix called P_{nxn} can be constructed (Ref 5:252-254).

$$P_{nxn} = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & \cdots & 0 & 0 \\ -\rho & 1 & 0 & \cdots & 0 & 0 \\ 0 & -\rho & 1 & \cdots & 0 & 0 \\ \vdots & & & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix}$$

where $|\rho|<1$. This matrix can be used to transform the original equation, Eq (7), into a new equation with independent error terms.

$$P_{nxn}Y_{nx1} = P_{nxn}X_{nxk}\beta_{kx1} + P_{nxn}\varepsilon_{nx1} \tag{9}$$

or for a model with a constant and a linear term

$$\underset{1}{\begin{bmatrix} \sqrt{1-\rho^2}Y_1 \\ Y_2-\rho Y_1 \\ \vdots \\ Y_n-\rho Y_{n-1} \end{bmatrix}} = \underset{2}{\begin{bmatrix} \sqrt{1-\rho^2} \\ 1-\rho \\ \vdots \\ 1-\rho \end{bmatrix}} \beta_0 + \underset{3}{\begin{bmatrix} \sqrt{1-\rho^2}X_1 \\ Y_2-\rho X_1 \\ \vdots \\ X_n-\rho X_{n-1} \end{bmatrix}} \beta_1 + \begin{bmatrix} \sqrt{1-\rho^2}\varepsilon_1 \\ \varepsilon_2-\rho\varepsilon_1 \\ \vdots \\ \varepsilon_n-\rho\varepsilon_{n-1} \end{bmatrix} \tag{10}$$

Applying Ordinary Least Squares to the transformed data, the data in the braces labeled 1, 2, 3, yields the best estimate of β_0 and β_1. This procedure is a special case of Generalized Least Squares. The problem, of course, is that ρ is not normally known. Various techniques have been invented to estimate both ρ and β_{kx1}. Four of these techniques will be discussed here: Cochrane-Orcutt, Durbin Nonlinear, Durbin Two Step and a method called Search.

Cochrane-Orcutt. One technique discussed by Cochrane and Orcutt, but not recommended by them because of possible bias, was to start with a guess of ρ in Eq (9) and then estimate β_{kx1} by Ordinary Least Squares (OLS) (Ref 9:53). Then the estimate of β_{kx1} was put back in the original equation, Eq (7), and the resulting residuals were used to estimate ρ. This new ρ was then put back in Eq (9) and a new β_{kx1} was estimated using OLS. The process was continued until successive estimates of ρ differed by a small amount. This method has been frequently illustrated without the first row of Eq (9). If the autocorrelation process has been going on before the first observation point, then leaving out the first row reduced the accuracy of this technique (Ref 10:96).

Durbin Nonlinear. Two of the other techniques discussed here that estimated ρ and β_{kx1} also do not use row 1 of Eq (10). One technique is called the Durbin Nonlinear (Ref 11). For the case of a constant plus a linear term, Durbin rewrites Eq (9) as

$$Y_t = \beta_0(1-\rho)+\beta_1 X_t - \rho\beta_1 X_{t-1} + \rho Y_{t-1} + \mu_t \tag{11}$$

or

$$Y_t = \beta_0' + \beta_1 X_t + \beta_2 X_{t-1} + \beta_3 Y_{t-1} + \mu_t \tag{12}$$

where μ_t is a random error. The Nonlinear technique minimizes the sum of squares of the errors in Eq (12) with the restriction that $\beta_2/\beta_1 - \beta_3$. β_3 will be the estimate of ρ. If desired, this ρ can be used in Eq (9) to estimate β_{kx1}. This technique, essentially least squares subject to a quadratic constraint, was tried with little success by Potluri and Griliches (Ref 11). This technique is difficult to program on a computer.

Durbin Two Step. The other technique discussed here that estimates ρ and β_{kx1} without row 1 is called the Durbin Two step (Ref 12:150-153). Durbin recommends directly performing least squares on Eq (12) while ignoring any constraint on the coefficients. The estimate of β_3 is then used as the estimate of ρ. If desired this ρ can be used in Eq (9) to estimate β_{kx1}. The Durbin Two step is certainly simple and is more effective than it first may appear. It leads to estimates having the same efficiency as the least square estimates in large samples (Ref 11; Ref 12:150-153).

Search. The last technique discussed under the assumption that $|\rho|<1$ is called the Search technique. Like the Durbin Nonlinear and the Durbin Two step it attempts to minimize the sum of squares of error of Eq (9), but unlike the Durbin techniques it includes row one of Eq (9). The Search technique is accurate. The minimum sum of squares is always found. Most important of all, the Search technique yields the maximum likelihood estimates of ρ and β_{kx1} (Ref 13: Ref 14). The Search technique is simple to understand. For successive increments of ρ between 1 and -1, excluding 0, OLS is performed on Eq (9) yielding estimates of ρ, and β_{kx1}, and the variance, S^2. There will be only one minimum point for S^2. ρ, and β_{kx1} corresponding to this minimum will be the maximum likelihood estimates (Ref 13).

The Search technique will probably require more iterations than Cochrane-Orcutt or the Durbin techniques; however, with careful planning the number of iterations can be reduced. If one knows the sign of ρ (ie., positive or negative autocorrelation) then only one half the interval of ρ need be searched. Since there is only one minimum S^2 point, not every increment of ρ need to evaluated. The interval of under search can be bisected and the midpoint and endpoints can be evaluated for S^2. The two resulting intervals can be bisected and their midpoints can be evaluated for S^2. The process continues and the minimum point of S^2 is located in successively smaller intervals. Knowing the sign of ρ and using the above process can reduce the maximum number of it-

erations needed to estimate ρ. No more than 14 iterations are necessary to estimate ρ to 2 decimal places.

Prediction. Goldberger showed how to predict future values of Y along with their prediction variances, σ_t^2 using ρ and β_{kx1} determined from the previous four techniques (Ref 15:372-374). The best linear unbiased estimate of future values of Y is given by

$$Y_t = X_{1xk,t}\,\beta_{kx1} + \rho^t\,(Y_n - X_n\beta_{kx1}) \qquad (13)$$

where $X_{1xk,t}$ are the predictor variables for the t^{th} period in the future, and X_n is the last row of X_{nxk}. Y_n is the last row of Y_{nxk}, and Y_t is the predicted value for the t^{th} period. The variance of this prediction, σ_t^2, is given by

$$\sigma_t^2 = (1 - \rho^{2t})\sigma^2 + (1 - \rho^2)\sigma^2[1 + (X_{1xk,t} - \rho X_n)'$$
$$(X'_{nxk}\,\rho'_{nxn}\,P_{nxn}\,X_{nxk})^{-1}(X_{1xk,t} - \rho X_n)] \qquad (14)$$

where P_{nxn} is as given in Eq (9), and $(1 - \rho^2)\sigma^2$ is the variance of the errors in Eq (9).

This concludes the discussion of the various techniques used if one decides to use the first order autoregressive model with the assumption that $|\rho|<1$. Next a technique used to estimate ρ for the assumption that $-\infty<\rho<\infty$ will be discussed.

$-\infty<\rho<\infty$ Assumption

The assumption that $-\infty<\rho<\infty$ is rarely discussed in the literature. ρ greater than one means the errors go to infinity in time. This is sometimes called an explosive pro-

cess. Although an explosive process may seem to be unlikely, some models may have errors that increase slightly with time.

Bayesian. Bayesian techniques have been used to develop the probability of ρ given the data for the assumption that $-\infty<\rho<\infty$ (Ref 16). The result is that the probability of ρ given the data, is inversely related to the variance of the error resulting from OLS performed on Eq (9) without row one. The smallest variance will correspond to the most likely estimate for ρ. Unfortunately, if this estimate of ρ is greater than +1 or less than -1, then the theory stops. No one has discussed how to predict future values if $|\rho|>1$. However, the Bayesian technique was used in this thesis to get an idea of how big ρ was when other estimates of ρ indicated ρ was greater than .99.

ρ = 1 Assumption

The simplest of all techniques used to correct for autocorrelation is also the most restrictive. It assumes $\rho = 1$. The technique is called first differences. For illustration assume a model with a constant and a linear term. Subtracting successive rows of Eq (7)

$$Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t \tag{15}$$

$$Y_{t-1} = \beta_0 + \beta_1 X_{t-1} + \varepsilon_{t-1} \tag{16}$$

along with $\rho = 1$ in Eq (8) yields

$$Y_t - Y_{t-1} = \beta_1 (X_t - X_{t-1}) + \mu_t \tag{17}$$

or

$$\Delta Y_t = \beta_1 \Delta X_t + \mu_t \tag{18}$$

Eq (18) has independent error terms μ_t. Ordinary Least Squares can be performed on Eq (18) yielding an estimate of β_1 of the original equation, Eq (15). If the assumption of the first order model is correct, and if ρ is close to 1, then the first difference estimate of β_1 will be better than the Ordinary Least Squares estimate. Also the corresponding first differences prediction of Y will be better than the Ordinary Least Squares estimate of Y (Ref 8).

This concludes the discussion of techniques used when a first order autoregressive model is assumed: Cochrane-Orcutt, Durbin Nonlinear, Durbin Two Step, Search, Bayesian, and First Differences. These techniques along with their advantages and disadvantages are summarized in Fig. 4.

<u>Use Per Cent Change Regression</u>

Using per cent change regression is not recommended in textbooks or journals as one of the procedures to use if autocorrelation is detected. However, it is mentioned here because ASD used it in their 110C report. See Chapter II. This technique is similar to first differences, but unlike first differences it really is a complete respecification of the model rather than a correction for autocorrelation. However, per cent changes may be a useful model when the errors of the original model were not only autocorrelated but their size also increased when the independent and dependent variables grew larger.

This concludes the discussion of techniques that can be used when autocorrelation is detected: ignore it, try a dif-

Assumption	Technique	Programming (1) simplest to (6) hardest	Computer Time To Execute (1) quickest to (6) longest	Quality of Estimates
$\rho = 1$	First Differences	(1) Simplest	(2) Two Regressions per index	Better than OLS when ρ is close to one.
$\lvert\rho\rvert<1$	Durbin 2 Step	(2) Simple	(2) Two Regressions per index	Best two step technique Better than OLS
	Cochrane-Orcutt Iterative	(5) Complex Tricky	(3) Less than Search	Technique not recommended by C-O themselves. Should be better than OLS
	Durbin Nonlinear	(6) Very complex	(4) Depends on iterative technique used	If not done properly, estimates will be worse than Durbin 2 Step. If done correctly should be better than OLS.
	Search	(4) Medium Easy to understand but code is long	(5) Usually less than 14 regressions Less than one minute of computer time per index used by author	Maximum likelihood estimates of ρ and β. Best estimates for large samples
$-\infty<\rho<\infty$	Bayesian	(3) Simple	(6) One regression per increment of ρ used. Depends on how high ρ goes	Useful only to get an estimate of ρ

Fig. 4 Techniques to use in presence of autocorrelation

ferent specification of the model, assume a first order auto-regressive model, or use per changes. The next chapter will investigate the formulation of the historical indices.

IV. Historical Indices

Preface

One basic goal of this thesis is to compare different techniques that may be used to predict six aeronautical price indices. Before the techniques were compared, it was thought that the historical price indices should be checked for validity. Invalid indices could effect the results of the thesis. The validity of the historical price indices depends upon the methodology used to calculate the indices, and the accuracy of the calculations. In this thesis the methodology used to calculate the historical price indices will be assumed to be correct. However, the calculations will not be assumed to be correct. Whenever feasible, the historical indices will be recalculated. This chapter consists mainly of examples of these recalculations. The recalculated price indices are listed in full in Appendix One. The recalculations were done not only to check the validity of the indices, but also to gain a better understanding of the innerworkings of the ASD forecasts.

Weights Used

The following weighted averages were used by ASD to calculate the six composite aeronautical price indices.

Airframe (AF)

(.4) AF Raw Material + (.225) Aircraft Parts & Equipment + (.375) 7.3% = Materials Overhead (19)

(.1425) Manufacturing Labor + (.2125) Engineering Labor + (.1924) Materials + (.4526) 7.8% Overhead = Development Airframe Index (20)

(.2010) Manufacturing Labor + (.0650) Engineering Labor + (.3040) Materials + (.4300) 7.8% Overhead = Production Airframe Index (21)

Engine

(.4) Engine Raw Material + (.6) Aircraft Engine Labor = Purchased Parts (22)

(.2701) Aircraft Engine Labor + (.3415) Engineering Labor + (.1942) Engine Raw Material + (.1942) Purchased Parts = Development Engine Index (23)

(.4184) Aircraft Engine Labor + (.0416) Engineering Labor + (.2700) Engine Raw Material + (.2700) Purchased Parts = Production Engine Index (24)

Avionics

$$(.10)\ \text{Avionics Raw Material} + (.35)\ \text{Development Avionics Labor} + (.55)\ \text{Development Overhead} = \text{Development Avionics Index} \qquad (25)$$

$$(.20)\ \text{Avionics Raw Material} + (.30)\ \text{Production Avionics Labor} + (.5)\ \text{Production Overhead} = \text{Production Avionics Index} \qquad (26)$$

The weights and methodology used were from ASD workshops, a Rand report, and a Navy study (Ref 4: Ref 17: Ref 18). Note, within each major area, airframe, engine, avionics, the production and development indices are closely related. Each of the subindices listed in Eqs (19) through (26) will now be discussed in length.

Airframe Raw Material Subindex

The airframe raw material subindex is a weighted average of the Wholesale Price Indexes (base year 1967) for 8 metals. The methodology is based on the 1970 Rand report (Ref 4). The Wholesale Price Index (WPI) for all 8 metals for years 1958 to 1975 was not readily available. Three sources were needed. Monthly issues of the Wholesale Price Index and Indexes periodical, a special computer listing of selected WPI's from the Department of Labor, and WPI's listed in the 1970 Rand report were all used (Ref 19: Ref 20: Ref 4).

Data from the three sources were checked where it overlapped to ensure consistency. The following is a list of the metals used, the Bureau of Labor Statistics codes, the weights used, and an example of the computations for the 1970 Airframe Materials Subindex:

Metal	Code	Weights	1970 WPI (1967=100)
Finished Steel	10-13-02	.02	114.4
Stainless Steel	10-13-02-53	.04	133.7
Titanium Sponge	10-22-01-56	.07	95.5
Aluminum Sheets	10-25-01-01	.29	110.7
Aluminum Rods	10-25-01-13	.11	93.4
Aluminum Extrusions	10-25-01-17	.20	120.6
Wire and Cable	10-26	.12	126.7
Nuts and Bolts	10-81	.15	121.9
1970 Airframe Raw Material Subindex			114.28

Aircraft Parts and Equipment Subindex

This index is actually a labor wage index. It is the average hourly earnings for aircraft parts and equipment production workers, Bureau of Labor Statistics code 3723,9 converted into a 1967 base. The index is developed by dividing the average hourly earnings for each year by the average hourly earnings in 1967 and then multiplying by 100. The data were obtained from Bureau of Labor Statistics Employment and Earnings publication (Ref 21). The following is an example calculation for the 1970 index:

Average hourly earnings for 1967, \$3.35

Average hourly earnings for 1970, \$3.99

1970 Aircraft Parts & Equipment Subindex =

$$\frac{3.99}{3.35} \times 100 = 119.10 \qquad (27)$$

Manufacturing Labor Subindex

This index is the average hourly earnings for aircraft production workers. Buréau of Labor Statistics code 3721, converted into a 1967 year base. The index is developed in the exact same way as the Aircraft Parts and Equipment subindex. The data is from the exact same place (Ref 21). The following is an example calculation:

Average hourly earnings for 1967, \$3.49

Average hourly earnings for 1970, \$4.17

1970 Manufacturing Labor Subindex:

$$\frac{4.17}{3.49} \times 100 = 119.48 \qquad (28)$$

Engineering Labor Subindex

This index is the average hourly earnings for engineers based on year 1967. The data were supplied to ASD from six large aerospace contractors. The 1974 and 1975 earnings were estimated. No example calculations are given because it might reveal proprietary information.

7.8% Overhead Subindex

This index is simply 7.8% compounded annually back and forth from a 1967 base of 100. The 7.8% was based on over-

head data provided to ASD from the same large aerospace contractors that supplied the engineering data. The following is an example calculation:

$$7.8\% \text{ Overhead Subindex for 1970} = (100)(1 + .078)^3$$
$$= 125.27 \qquad (29)$$

7.3% Overhead Subindex

This index is simply 7.3% compounded back and forth from a 1967 base of 100. Mr. Benoy of ASD stated the 7.3% figure was based on vendors having lower indirect labor costs than prime contractors (Ref 22). The following is an example calculation:

$$7.3\% \text{ Overhead Subindex for 1964} = (100)(1 + .073)^{-3}$$
$$= 80.95 \qquad (30)$$

Engine Raw Material Subindex

This index is a weighted average of prices for 8 metals converted into a 1967 base. The prices were determined by averaging weekly prices of the metals listed in <u>Iron Age</u> magazine. The metals selected and the weights used were based upon ASD analysis of data from United States Air Force Aircraft Systems Program Offices. Calculation of this index is a very long process. Thus only the 1975 figures were partially checked. The rest of the data were assumed to be correct as listed in ASD's 110C report (Ref 3). The following is a list of the metals used, the weights used, and the 1975

average prices:

Metal	Weights	1975 Prices (¢/lb)
Titanium	.329	225.0
Chromium	.165	235.9
Aluminum	.092	39.6
Cobalt	.041	397.5
Molydenum	.030	259.9
Vanadium	.009	255.1
Magnesium Ingot	.007	81.8
Nickel	.327	204.5
	1975 weighted average	210.43

Engine Raw Material Subindex for 1975
= (1975 prices/1967 prices) × 100
= (210.43/105.12) × 100
= 200.18 (31)

Aircraft Engine Labor Subindex

This index is the average hourly earnings for aircraft engine and engine parts production workers, Bureau of Labor Statistics code 3722. The index is developed by dividing the average hourly earnings for each year by the average hourly earnings in 1967 and then multiplying by 100. The data were obtained from the Bureau of Labor Statistic Employment and Earnings publication (Ref 21). The following is an example calculation:

Average hourly earnings for 1967, \$3.42

Average hourly earnings for 1970, \$4.10

1970 Aircraft Enginer Labor Subindex =

$$\frac{4.10}{3.42} \times 100 = 119.88 \qquad (32)$$

Avionics Raw Material Subindex

All the Avionics subindices are based on methodology developed in a 1970 United States Navy report (Ref 17). This index is simply the arithmetic average of two Wholesale Price Indexes (WPI), base year 1967. WPI 1172, Integrating and Measuring Instruments, and WPI 1178, Electronic Components and Accessories. The following is an example calculation:

$$\text{1975 Avionics Raw Material Subindex} = (.5)(1172\ \text{WPI}) + (.5)(1175\ \text{WPI}) = .5(140.0) + .5(115.5) = 127.8 \quad (33)$$

Development Avionics Labor Subindex

This index is a weighted average of hourly wages for engineers, engineering technicians, draftsmen, and production workers, converted into a 1967 base. ASD used the wage data supplied or estimated in the 1974 Navy report, except for 1974 and 1975 where they estimated the engineers' wages themselves (Ref 18). The following is an example calculation:

$$\text{1975 weighted wages} = .5(\text{engineer ave hrly wage}) + .2(\text{engineering technician}) + .15(\text{draftsmen}) + .15(\text{production workers}) = .5(10.78) + .2(6.3) + .15(6.1) + .15(4.66) = 8.264 \quad (34)$$

$$\text{1975 Development Avionics Labor Subindex} = (\text{1975 wages}/\text{1976 wages}) \times (100) = (8.264/5.0975) \times (100) = 162.12 \quad (35)$$

Production Avionics Labor Subindex

This index uses the same wages the Development Avionics Labor subindex, but uses different weights. The following is an example calculation:

$$1975 \text{ weighted wages} = .2(\text{engineers}) = .15(\text{engineering technicians}) + .05(\text{draftsmen}) + .6(\text{production workers}) = .2(10.78) + .15(6.3) + .05(6.1) = .6(4.66) = 6.202 \quad (36)$$

1975 Production Avionics Labor Subindex:

$$(1975 \text{ wages}/1967 \text{ wages}) \times (100) = (6.202/3.804) \times (100) = 163.04 \quad (37)$$

Development Avionics Overhead Subindex

This index is simply the Development Avionics Labor subindex times 1.3274% compounded annually back and forth from 1967. Thus overhead as a percentage of Avionics Labor was simply assumed to increase each year by 1.3274%. The methodology is taken from the 1970 Navy report (Ref 17). The following is an example calculation:

$$1975 \text{ Development Avionics Overhead Subindex} = 162.12\ (1 + .013274)^8 = 180.16 \quad (38)$$

Production Avionics Overhead Subindex

This index is simply the Production Avionics Labor subindex times 1.3274% compounded annually back and forth from

1967. Again overhead as a percentage of Avionics Labor was assumed to increase each year by 1.3274%. The methodology is again from the Navy study (Ref 17). The following is an example calculation:

$$\text{1975 Production Avionics Overhead Subindex} = 163.04\ (1 + .013274)^{8} = 181.18 \qquad (39)$$

That concludes a brief explanation of the methods used in the preparation of the historical values. The recalculated indices are listed in Appendix A. The recalculated indices compare very closely to the ASD calculations in the 110C report with one exception (Ref 4). The ASD Airframe Raw Material Index value was slightly lower than the recalculated values prior to 1967. This was due to differences in the ASD calculations when the Bureau of Labor Statistics switched from a 1957-1958 base to a 1967 base. ASD has recalculated the Airframe Raw Material Subindex values for use in future reports.

One can see that there are basically three sets of indices: labor wage indices, raw material price indices, and overhead cost indices. The labor wage indices were based on average hourly earnings. The raw material indices were a combination of Wholesale Price Indexes, and the overhead indices increased at a given percentage each year. Many weighted averages were used to construct the historical indices. It is clear that one could spend the entire time in a thesis discussing the choice and development of price indices. In

fact, a 1971 Air Force Institute of Technology thesis did just that (Ref 23). However, it was decided not to investigate the historical indices any further. The rest of the thesis will be spent investigating the prediction of the constructed indices. First, however, the independent/predictor variables will be discussed.

Independent/Predictor Variables

There are five variables in this thesis that were used or tried as independent/predictor variables in the regressions. These variables are: GNP price deflator, Primary Metals Index, Electrical Machinery Index, the money supply M1, and the U.S. total unemployment rate. The aircraft industry unemployment rate would have been used instead of the total U.S. unemployment rate, but data were only available on the aircraft unemployment rate after 1964 (Ref 24). Historical values for the Primary Metals Index and the Electric Machinery Index were taken from the ASD 110C report (Ref 4). The original source was Wharton Econometric Forecasting Associates, Inc. Historical values for the GNP price deflator, M1, and the national unemployment rate were taken from the January 1976 Economic Report of the President (Ref 25). The historical independent variables are all listed in Appendix B. Predictions for some of these variables are also listed in Appendix B. The predictions are from the Wharton Econometric Forecasting Associates, Inc. More will be said about these predictions in Chapter VII.

This concludes a discussion of the historical data. The next chapter will compare a technique used in the 110C report, subindex per cent change regression, with a similar technique, aggregate per cent change regression.

V. Per Cent Change Regressions

Preface

This chapter describes two distinct approaches to index prediction using per cent change regression. The first approach is the subindex method used by ASD in the 110C report. See Fig. 1. The regression equations using this approach are shown below along with R^2, the coefficient of determination, the t statistics for the coefficients, DW, the Durbin Watson statistic for autocorrelation, and S^2, the calculated variance. Predictions and prediction intervals will also be shown.

The second approach involves first aggregating the data, and then performing the regressions. The regression equations for this second approach will be shown along with R^2, the t statistics, DW, and S^2. The two approaches are compared in the end of the chapter. All of the calculations in this chapter were performed on the indices recalculated in Chapter IV.

Subindex Method

Regression Equations

$$\Delta \text{ \% Engineering Labor} = \underset{(1.76)}{1.582} + \underset{(4.27)}{.8286}\ \Delta \text{ \% } GNP_d$$

$$DW = 2.01 \qquad R^2 = .55 \qquad S^2 = 3.8739 \qquad (40)$$

$$\Delta\ \%\ \text{Manufacturing Labor} = \underset{(2.99)}{2.124} + \underset{(5.60)}{.8602}\ \Delta\ \%\ GNP_d$$

$$DW = 2.40 \qquad R^2 = .68 \qquad S^2 = 2.4342 \tag{41}$$

$$\Delta\ \%\ \text{AC Engine Labor} = \underset{(3.63)}{2.281} + \underset{(5.71)}{.7748}\ \Delta\ \%\ GNP_d$$

$$DW = 2.15 \qquad R^2 = .68 \qquad S^2 = 1.8784 \tag{42}$$

$$\Delta\ \%\ \text{AC Parts \& Equipment} = \underset{(5.40)}{2.321} + \underset{(7.19)}{.6673}\ \Delta\ \%\ GNP_d$$

$$DW = 2.37 \qquad R^2 = .77 \qquad S^2 = .8888 \tag{43}$$

$$\Delta\ \%\ \text{Development Avionics Labor} = \underset{(6.40)}{3.445} + \underset{(3.81)}{.4429}\ \Delta\ \%\ GNP_d$$

$$DW = 1.78 \qquad R^2 = .49 \qquad S^2 = 1.3955 \tag{44}$$

$$\Delta\ \%\ \text{Avionics Production Labor} = \underset{(5.35)}{2.914} + \underset{(4.55)}{.5361}\ \Delta\ \%\ GNP_d$$

$$DW = 2.44 \qquad R^2 = .58 \qquad S^2 = 1.4281 \tag{45}$$

$$\Delta\ \%\ \text{AF Raw Material} = \underset{(-.87)}{-1.381} + \underset{(4.46)}{1.143}\ \Delta\ \%\ \text{P Metal}$$

$$DW = 1.98 \qquad R^2 = .57 \qquad S^2 = 24.461 \tag{46}$$

$$\Delta\ \%\ \text{Engine Raw Material} = \underset{(-.96)}{-1.516} + \underset{(5.05)}{1.296}\ \Delta\ \%\ \text{P Metal}$$

$$DW = 1.82 \qquad R^2 = .63 \qquad S^2 = 24.565 \tag{47}$$

$$\Delta\ \%\ \text{Avionics Raw Material} = \underset{(1.72)}{.7964} + \underset{(4.93)}{.6907}\ \Delta\ \%\ \text{E Mach}$$

$$DW = 2.25 \qquad R^2 = .62 \qquad S^2 = 3.5308 \tag{48}$$

where $\Delta\ \%$ is $\dfrac{\text{Index}_t - \text{Index}_{t-1}}{\text{Index}_{t-1}} \times 100$

GNP_d is GNP Price Deflator

PMetal is Primary Metals Index

E Mach is Electrical Machinery Index

Numbers in parenthesis are t statistics for the coefficients

All equations except that for Airframe Raw Material are very similar to the equations listed in ASD report 110C (Ref 3: C-1). The difference in the Airframe Raw Material equations was a direct result of differences in calculation of the historical data. See Chapter IV. The corresponding equation listed in report 110C is:

$$\Delta \text{ \% AF Raw Material} = 1.289 + 1.146 \, \Delta \text{ \% PMetal} \qquad (49)$$

For all regression equations, visual residual analysis did not indicate any strong departure from the assumptions of Ordinary Least Squares. The Durbin Watson statistics all tend to indicate that the errors were not autocorrelated. However, four coefficients were found not to be significantly different from Ø at the 95% confidence level. The four coefficients were the constants in each of the raw material regressions and the constant in the engineering labor regression. The regression equations will be used, as is, with the constant terms present, to predict future values. This is done so that a direct comparison can be made with the 110C methodology

Predictions

Changing per cent change predictions into subindex pre-

dictions is a simple process. The per cent change predictions are repeatedly applied to the last known index, the 1975 value. The following is an example of the equations used:

$$76 \text{ index} = 75 \text{ index } \left(1 + \frac{\Delta\%_{76}}{100}\right) \tag{50}$$

$$77 \text{ index} = 75 \text{ index } \left(1 + \frac{\Delta\%_{76}}{100}\right)\left(1 + \frac{\Delta\%_{77}}{100}\right) \tag{51}$$

where $\Delta\%_{76}$ stands for the per cent change prediction for 1976. The predictions for the subindices are listed in Table I. The predictions for all subindices are almost identical to their counterparts in the 110C report (Ref 3:12-14). After the subindices are predicted, they are weighted and aggregated using Eqs (19) through (25). This yields the predictions for the final six indices. These predictions are listed in Table II. No direct comparison of these final six indices can be made with the 110C report. The final six indices are listed in the 110C report only in a Fiscal Year base.

Prediction Intervals

Only 95% prediction intervals will be given in this thesis. If the regression model is correct and continues to hold into the future, and if the Wharton predictors are known without error, then the following could be said about the prediction intervals. If one could somehow collect new historical data on the same indices, do the regressions again,

Table I

Subindex Predictions and Prediction Intervals: Using Per Cent Change Regressions

CY	Engineering Labor	Manufacturing Labor	AC Parts & Equipment
1976	182.97± 7.09	190.48±5.81	175.12±3.25
1977	195.17±10.39	204.59±7.00	186.36±4.86
1978	210.10	221.81	199.79
1979	227.09	241.50	214.89
1980	242.98	260.21	229.25
1981	256.51	276.50	241.93
1982	273.39	296.72	257.29
1983	290.94	317.92	273.28
1984	304.68	335.03	286.53
1985	316.20	349.81	298.26

CY	AC Engine Labor	Development Avionics Labor	Production Avionics Labor
1976	188.44±5.09	171.96±4.01	172.97±4.08
1977	201.70±7.48	182.56±5.87	183.71±5.97
1978	217.74	194.77	196.27
1979	235.95	208.26	210.25
1980	253.28	221.46	223.74
1981	268.49	233.81	236.03
1982	287.16	248.11	250.54
1983	306.69	263.07	265.68
1984	322.69	276.55	278.82
1985	336.69	289.33	290.92

CY	AF Raw Material	Engine Raw Material	Avionics Raw Material
1976	175.62±17.00	213.55±20.55	135.90±5.62
1977	188.31±24.91	231.15±30.26	143.89±8.12
1978	201.77	250.00	151.78
1979	215.04	268.76	159.60
1980	228.80	288.40	167.35
1981	239.66	304.07	175.06
1982	254.56	325.66	182.73
1983	269.38	347.32	190.38
1984	278.67	361.06	198.00
1985	284.66	370.04	205.62

Table II

Composite Index Predictions and Prediction Intervals: Using Subindex Per Cent Change Regressions

CY	Development Airframe	Production Airframe
1976	189.70±1.92	189.54±1.83
1977	203.70±2.81	203.64±2.69
1978	219.49	219.43
1979	236.79	236.60
1980	254.44	254.25
1981	271.71	271.58
1982	291.50	291.43
1983	312.49	312.47
1984	332.50	332.56
1985	352.43	352.59

CY	Development Engine	Production Engine
1976	193.40±5.89	197.70± 7.13
1977	207.48±8.66	212.56±10.50
1978	223.90	229.62
1979	241.85	247.98
1980	259.31	266.13
1981	274.07	281.44
1982	292.93	301.14
1983	312.36	321.39
1984	326.97	336.45
1985	338.76	348.44

CY	Development Avionics	Production Avionics
1976	180.27±3.79	176.46±4.00
1977	192.84±5.59	188.69±5.87
1978	207.19	202.69
1979	223.03	218.14
1980	238.83	233.38
1981	247.39	247.77
1982	271.42	264.38
1983	289.79	281.83
1984	306.92	297.69
1985	323.59	312.83

and each time calculate 95% predictions intervals, then the probability is that 95 out of 100 of the prediction intervals calculated would contain the actual future value of the index (Ref 25:304). However, one cannot collect the historical data again. In practice, then, the following is said: There is a 95% probability that the actual future index values will fall within the prediction intervals. This of course, assumes that the relationship between the predictor variables and the indices will continue to be the same, and that the Wharton Predictions are without error. The effect of errors in the Wharton predictors will be discussed in Chapter VII.

Calculation of prediction intervals is a difficult process. In fact the calculations required estimates of the covariances between the subindices and the assumption of a near normal distribution for the predictions. Since the calculations were burdensome, only prediction intervals for 1976 and 1977 were calculated. The prediction intervals are listed in Tables I and II. Prediction intervals for future years will grow progressively larger and larger at a rate comparable to the aggregate method (See Table III). The full explanation of predicion interval construction is contained in Appendix C.

Aggregate Method

Regression Equations

$$\Delta\ \%\ \text{Air Frame Development} = \underset{(13.98)}{3.959} + \underset{(9.06)}{.5544}\ \Delta\ \%\ GNP_d$$

DW = 2.36 $R^2 = .845$ $S^2 = .3858$ (52)

$$\Delta\ \%\ \text{Air Frame Production} = \underset{(15.60)}{3.753} + \underset{(11.54)}{.6020}\ \Delta\ \%\ GNP_d$$

DW = 2.80 $R^2 = .900$ $S^2 = .2785$ (53)

$$\Delta\ \%\ \text{Engine Development} = \underset{(-1.70)}{-.9228} + \underset{(11.97)}{1.4054}\ \Delta\ \%\ GNP_d$$

DW = 2.42 $R^2 = .905$ $S^2 = 1.4210$ (54)

$$\Delta\ \%\ \text{Engine Production} = \underset{(-3.44)}{-1.825} + \underset{(14.13)}{1.6207}\ \Delta\ \%\ GNP_d$$

DW = 2.17 $R^2 = .930$ $S^2 = 1.3549$ (55)

$$\Delta\ \%\ \text{Avionics Development} = \underset{(6.52)}{3.3155} + \underset{(5.03)}{.5521}\ \Delta\ \%\ GNP_d$$

DW = 1.74 $R^2 = .627$ $S^2 = 1.2435$ (56)

$$\Delta\ \%\ \text{Avionics Production} = \underset{(4.37)}{2.0740} + \underset{(6.71)}{.7087}\ \Delta\ \%\ GNP_d$$

DW = 2.28 $R^2 = .761$ $S^2 = 1.0850$ (57)

where all symbols have the same meaning as used in the sub-index method

For all regression equations, visual residual analysis did not indicate any strong departure from the assumptions of Ordinary Least Squares. All the Durbin Watson statistics, except for Production Airframe, tended to indicate that the errors were not autocorrelated. The Durbin Watson statistic

for the Airframe regression was in the inconclusive region. Only one coefficient, the constant term for the Development Engine equation, was not significantly different from zero at the 95% confidence level. ASD did not use the aggregate per cent change method in any of their reports.

Predictions

The predictions were done using Eqs (50) and (51). No other steps are needed since the data has already been aggregated. The predictions are listed in Table III.

Prediction Intervals

Unlike the subindex method, the calculation of the prediction intervals did not require estimates of covariances; however, a near normal distribution was assumed for the predictions. The prediction intervals for all six composite indices are listed in Table III. The prediction interval for Production Avionics is shown in Fig. 5. The explanation of the prediction interval construction for the aggregate method is shown in Appendix C.

Comparison of the Aggregate and the Subindex Methods

It is important to realize that the difference between the subindex per cent change regression and the aggregate per cent regression method is not just the aggregation. There are differences in the independent variables and in the treatment of overhead between the two methods. The subindex method uses the Primary Metals Index as the independent variable for

Table III

Composite Index Predictions and Prediction Intervals: Using Aggregate Per Cent Change Regressions

CY	Development Airframe	Production Airframe
1976	189.77± 2.30	189.71±1.95
1977	203.74± 3.39	203.85±2.88
1978	220.08± 4.37	220.48±3.71
1979	238.37± 5.31	239.18±4.52
1980	256.45± 6.18	257.57±5.26
1981	273.43± 7.02	274.69±5.98
1982	293.41± 7.89	294.98±6.72
1983	314.52± 8.78	316.41±7.48
1984	333.58± 9.67	335.51±8.25
1985	351.70±10.60	353.46±9.04

CY	Development Engine	Production Engine
1976	194.70± 4.53	199.56± 4.52
1977	209.71± 6.67	215.78± 6.66
1978	229.34± 8.61	235.43± 8.63
1979	252.53±10.54	263.31±10.60
1980	273.40±12.35	286.41±12.47
1981	289.35±14.09	303.48±14.26
1982	311.21±15.88	327.62±16.12
1983	333.86±17.71	352.62±18.03
1984	348.56±19.53	307.85±19.93
1985	358.36±21.35	376.97±21.82

CY	Development Avionics	Production Avionics
1976	180.49± 3.96	175.42± 3.60
1977	192.59± 5.80	186.69± 5.27
1978	206.76± 7.45	200.25± 6.76
1979	222.58± 9.11	215.54± 8.18
1980	237.99±10.54	230.00± 9.49
1981	252.20±11.90	242.60±10.71
1982	268.96±13.31	258.01±11.98
1983	283.85±14.73	274.03±13.25
1984	299.20±16.13	287.08±14.51
1985	313.51±17.57	298.43±15.79

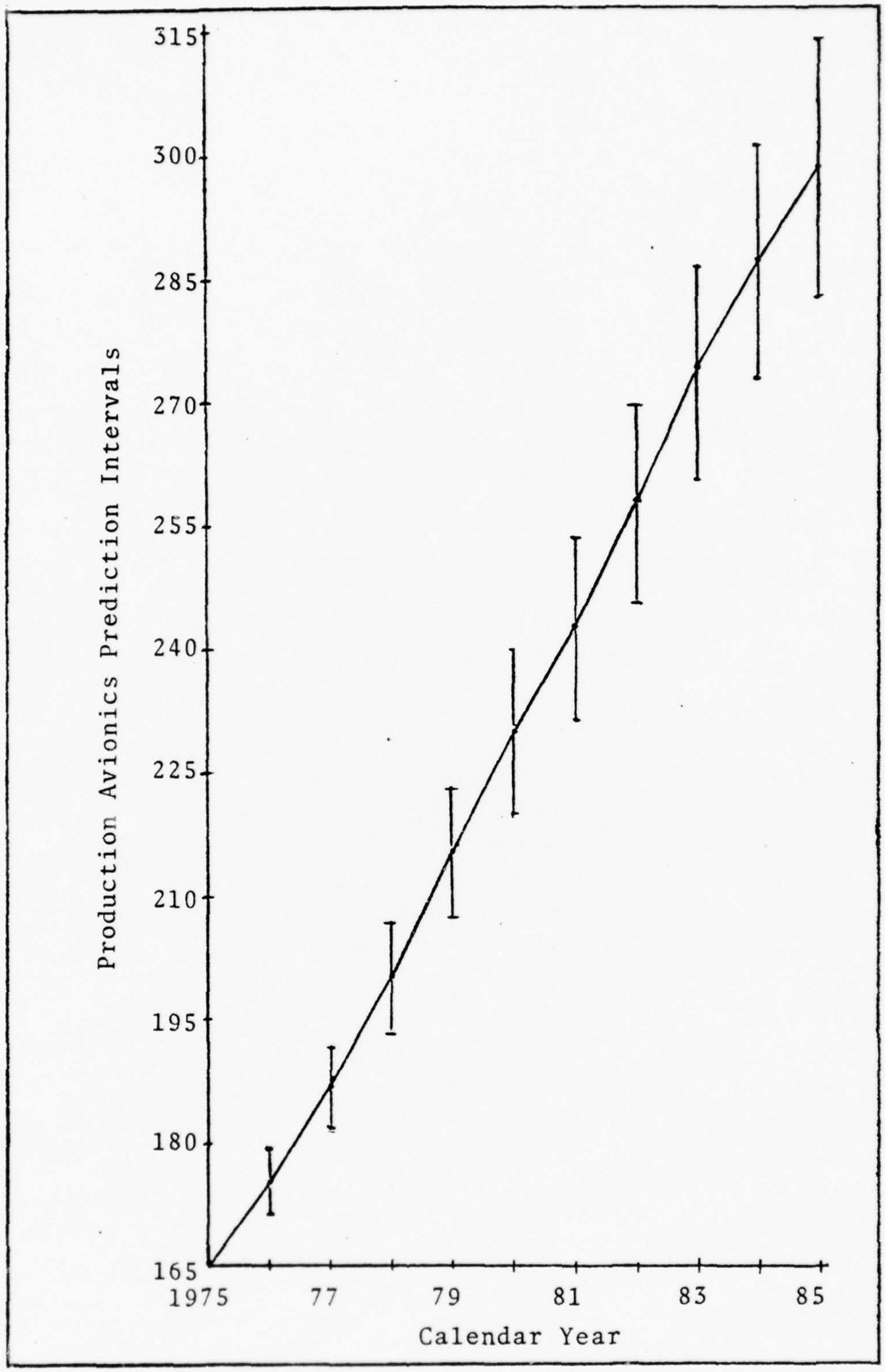


Fig. 5 Production Avionics Prediction Intervals vs. Calendar Year

the airframe and engine material regressions, and the Electric Machinery Index as the independent variable in the avionics materials regression. The GNP price deflator is used in the other subindex regressions. In the aggregate method, only the GNP price deflator is used as the independent variable in the regressions. Primary Metals and Electrical Machinery were tried as additional independent variables in the aggregate method, but they were found not to be significant at the 95% confidence level.

In the subindex method, overhead is not regressed but rather projected at a predetermined rate. For the airframe indices two different overhead subindices were used. One overhead subindex is 7.3% of the value for the previous year. The other overhead subindex is 7.8% of the value of the previous year. The avionics overhead as a percentage of avionics labor increases each year by 1.3274%. In the subindex method these overheads are projected into the future at their respective rates. In the aggregate method the overhead subindices are first aggregated with the other subindices and then regressed, and then predicted. See Fig. 1.

Keeping the procedural differences in mind, the predictions for both the subindex and aggregate method can be analyzed. The predictions of the airframe indices differed little between the methods. The aggregate method engine indices had slightly higher values than their subindex counterparts. The predictions for the avionics indices were significantly lower using the aggregate method than the subindex

method. The reason for the differences in the engine indices is not known. Part of the difference in the avionics indices is probably due to the dramatic difference in overhead calculation between the two methods.

Both methods have their own advantages and disadvantages. The aggregate method is computationally simpler and quicker. Only six regressions are done instead of nine. The weights used in report 110C are only estimates at best. The subindex method would permit a user of the report to use his own weights to develop production and development composite indices without having to do any further regressions. In the aggregate method a user who wants to use his own weights has to do the regressions over. However, as noted below, this probably presents no significant disadvantage. The prediction intervals give little advantage to one method over the other. Three indices have smaller prediction intervals for the aggregate method, and three indices have smaller prediction intervals for the subindex method.

Given three possible independent variables, the GNP price deflator, the Primary Metals Index, and the Electrical Machinery Index, and given the choice between the per cent change subindex method and the aggregate per cent change method, the author recommends the aggregate method. Computationally it is simpler and quicker. In addition, if a user were inclined to use his own weights he would probably want to make other changes, and might even do an entire report himself. Furthermore, the author prefers the treatment of overhead used in the aggregate method.

VI. Improving the Estimates

This chapter investigates different techniques that might be used to imporove ASD's estimation of the six composite aeronautical price indices. Techniques including simple regression, regression with nonlinear variables, multivariate regression, Generalized Least Squares, First Differences, and Per Cent Changes will be discussed.

Defining the Beginning

Before the investigation began, it was necessary to define a starting point. The historical subindices could have been used as the starting point or the subindex data could have been aggregated and the composite indices could have been used as the starting point. Unfortunately, there was not enough time to do parallel investigations, one for each possible starting point. The composite indices were chosen as the starting point for three reasons: first, fewer regressions were needed; second, ASD was contemplating using the composite technique in their next report; and third, the composite technique was recommended in Chapter V when dealing with per cent change regressions. It is important to realize that the results of this chapter are based upon starting with the composite data. Different relations that might result are discussed in the end of this chapter.

Simple Regression

Regressing the six composite indices against the GNP

price deflator is the easiest technique to use. The results are listed below.

$$\text{Development Air Frame} = \underset{(-10.22)}{-54.789} + \underset{(30.60)}{1.900}\ GNP_d$$

$$DW = .25 \qquad R^2 = .98 \qquad S^2 = 21.02 \tag{58}$$

$$\text{Production AF} = \underset{(-10.67)}{-53.743} + \underset{(32.30)}{1.884}\ GNP_d$$

$$DW = .24 \qquad R^2 = .98 \qquad S^2 = 18.54 \tag{59}$$

$$\text{Development Engine} = \underset{(-8.24)}{-20.532} + \underset{(153.93)}{1.558}\ GNP_d$$

$$DW = .65 \qquad R^2 = .99 \qquad S^2 = 4.55 \tag{60}$$

$$\text{Production Engine} = \underset{(-4.27)}{-18.363} + \underset{(30.91)}{1.542}\ GNP_d$$

$$DW = .42 \qquad R^2 = .98 \qquad S^2 = 13.56 \tag{61}$$

$$\text{Development Avionics} = \underset{(-7.80)}{-37.050} + \underset{(28.46)}{1.684}\ GNP_d$$

$$DW = .28 \qquad R^2 = .98 \qquad S^2 = 18.85 \tag{62}$$

$$\text{Production Avionics} = \underset{(-7.05)}{-22.855} + \underset{(40.49)}{1.521}\ GNP_d$$

$$DW = .39 \qquad R^2 = .99 \qquad S^2 = 7.69 \tag{63}$$

This was basically the same technique as used by ASD in the 110A and 110B reports. See Chapter II. The result is that the Durbin Watson statistics all indicate that the data may be highly autocorrelated. This is a warning that the estimates obtained using this technique may not be the best available.

To get a better idea of what was actually happening, the historical data were plotted along with the regression lines in Figs. 6 to 11. Careful examination of these plots yields interesting results. There is a definite pattern to the data. In Figs. 6, 7, 10, 11 the data starts below the regression line, remains there for awhile, and then goes above the line, and finally ends up below the line again. In Figs. 8 and 9 the exact opposite happens. From these patterns one can immediately tell that the Durbin Watson statistic will be very low indicating possible strong autocorrelation. The data would have to be displaced randomly above and below the line for the Durbin Watson statistic to indicate that the data were not autocorrelated. Thus, even though R^2 is very high and the data points are close to the regression line, the low Durbin Watson statistic indicates that procedures other than simple regression may be more accurate. Since accuracy is so important in the price index forecasting, other regression techniques will be investigated.

Regression with Nonlinear Variables

Would a nonlinear model fit the data better? A quadratic model would certainly fit well over a few years, but then the data swings back the other way and the quadratic model would no longer fit well. High order curves may fit the data well, but it would be hard to justify using a higher order curve. First, and most important of all, a linear

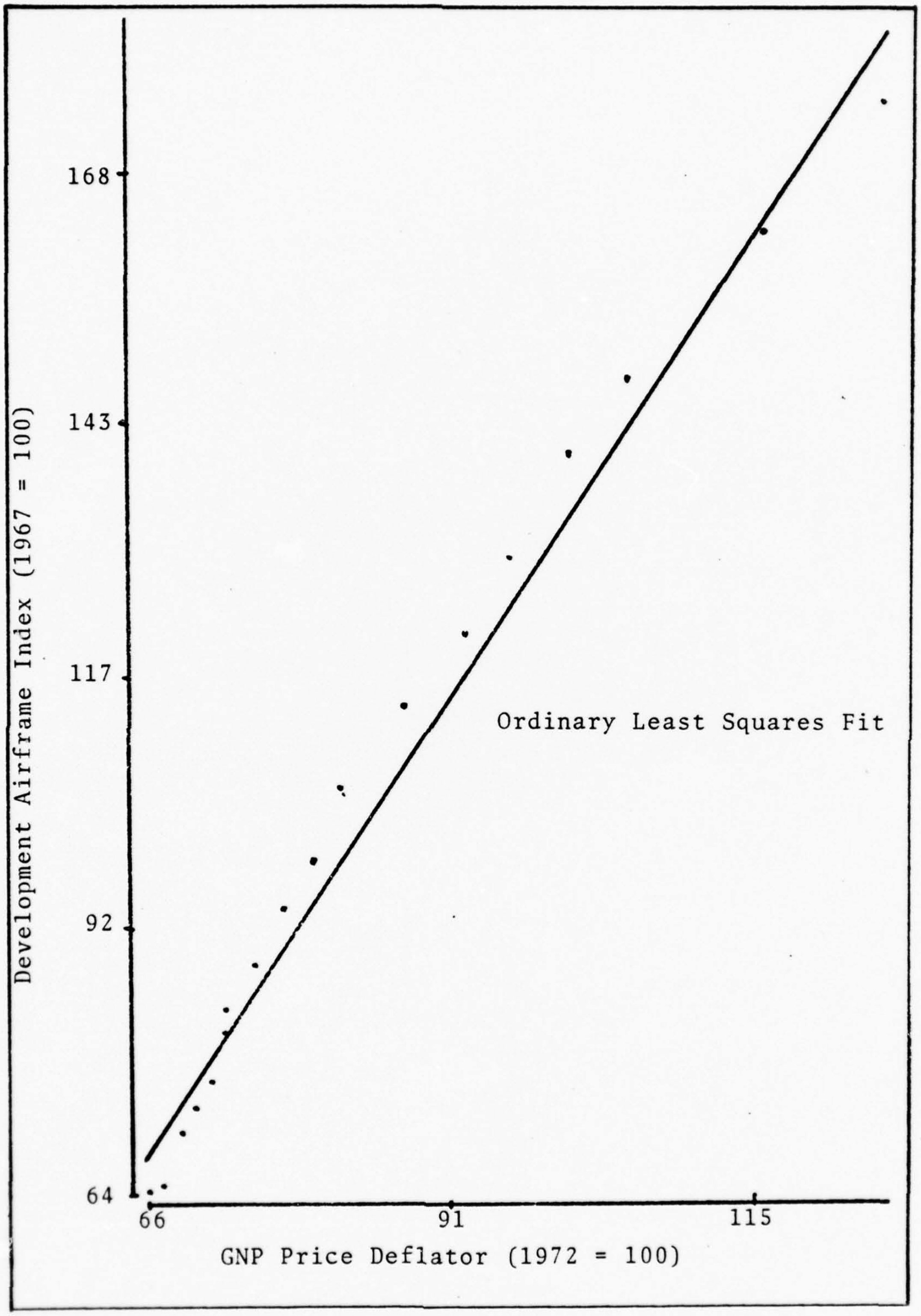


Fig. 6. Development Airframe Index vs. GNP Price Deflator

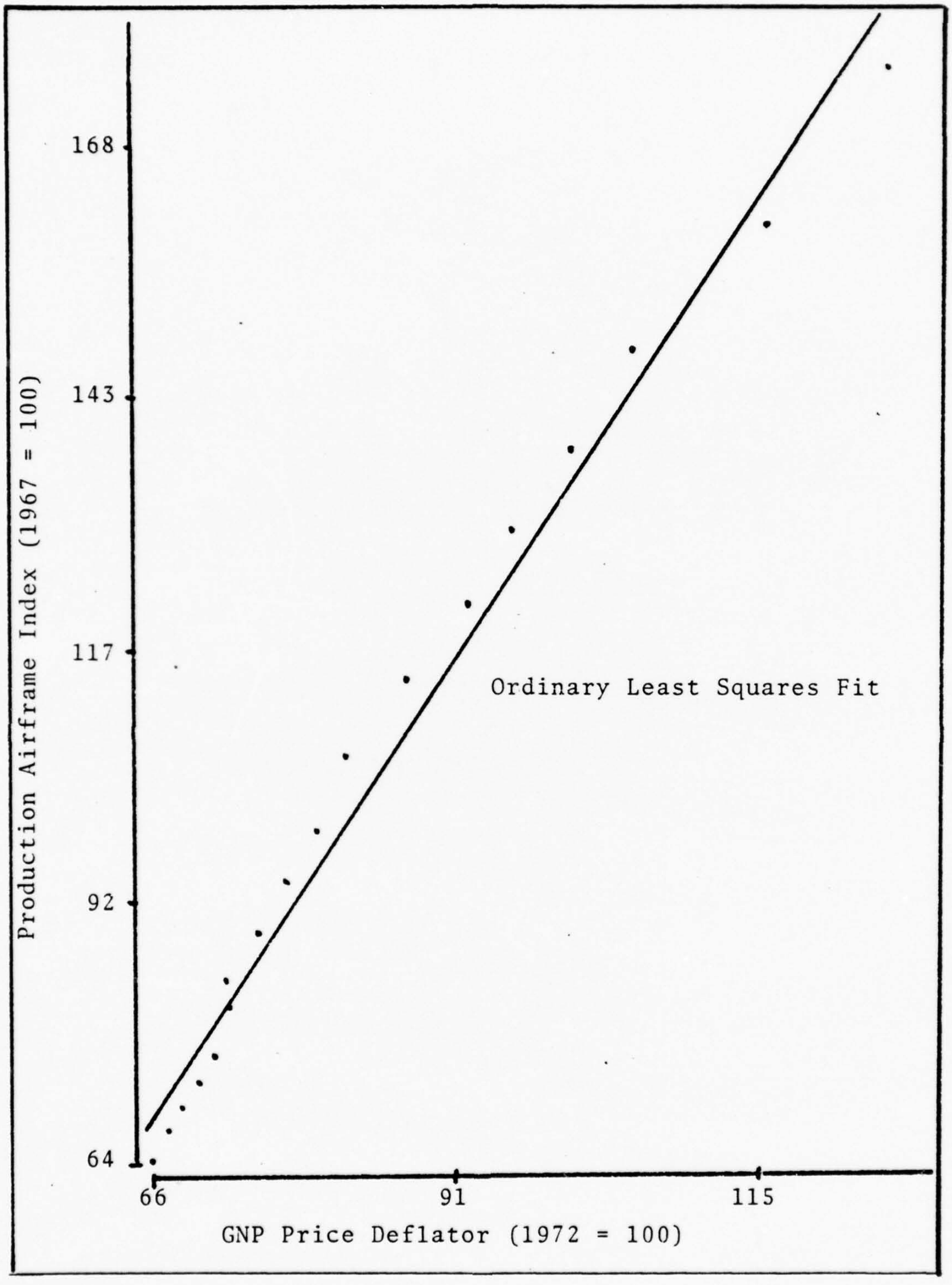


Fig. 7. Production Airframe Index vs. GNP Price Deflator

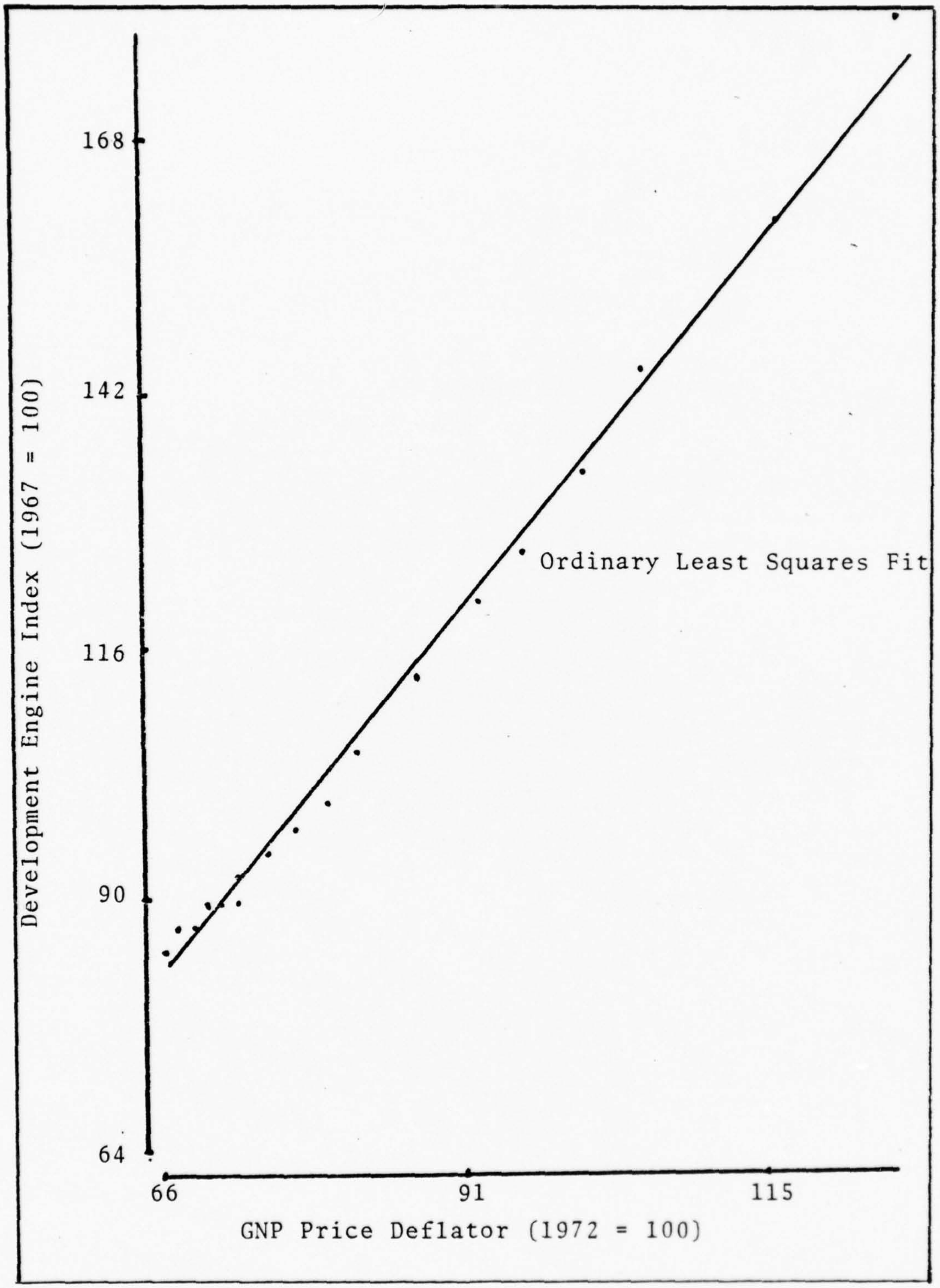


Fig. 8. Development Engine Index vs. GNP Price Deflator

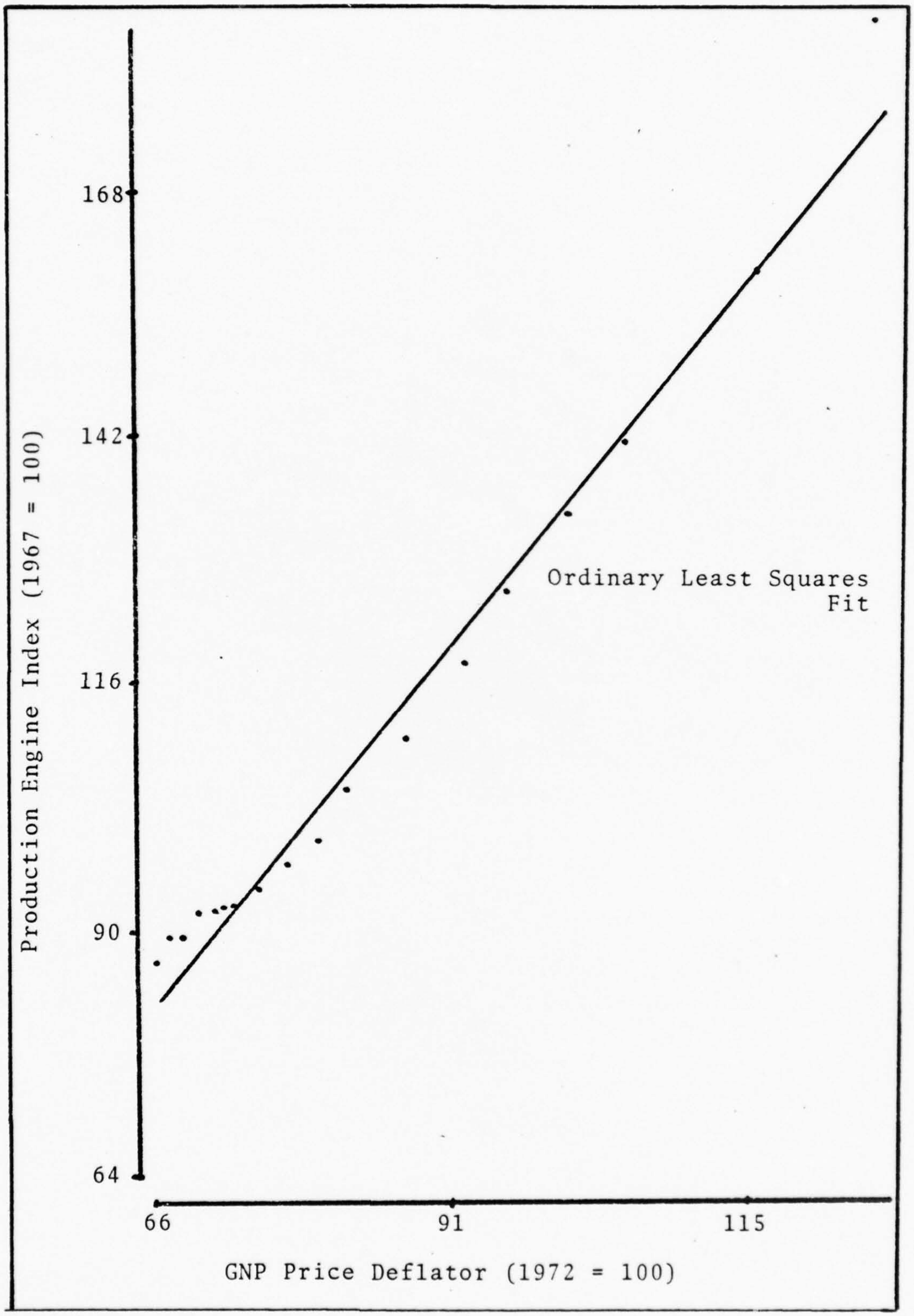


Fig. 9. Production Engine Index vs. GNP Price Deflator

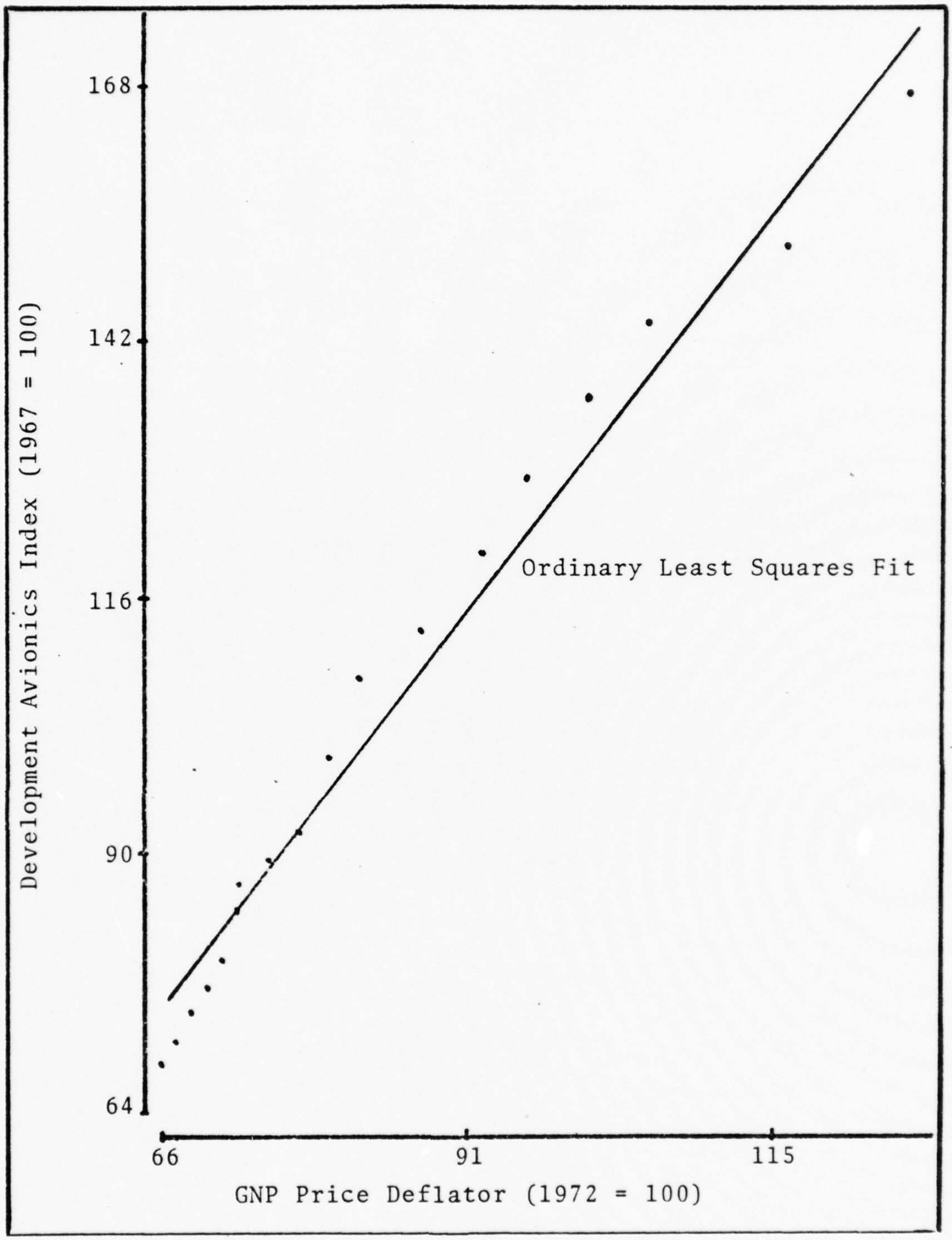


Fig. 10. Development Avionics Index vs. GNP Price Deflator

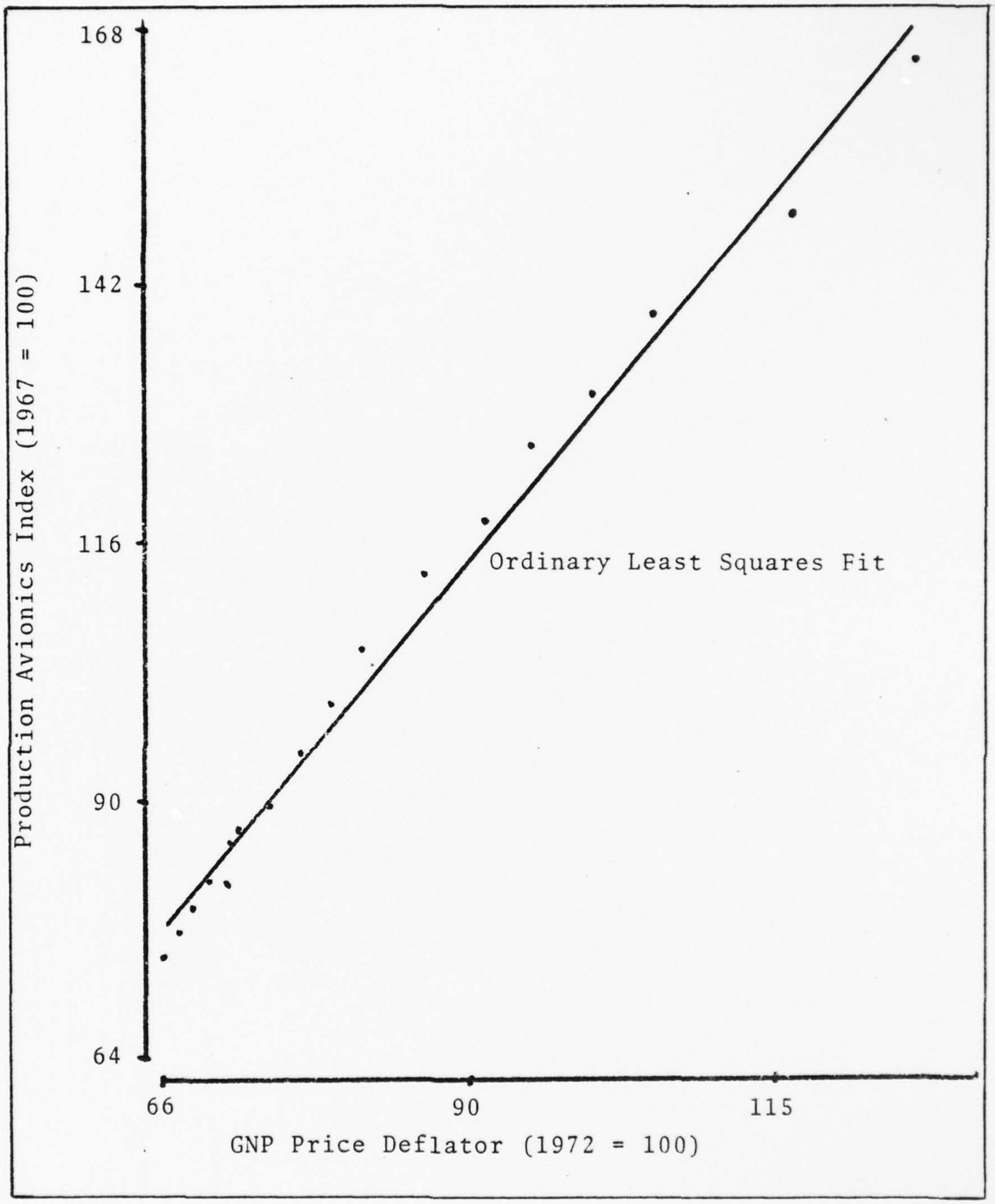


Fig. 11. Production Avionics Index vs. GNP Price Deflator

relationship between the GNP price deflator and the composite indices makes economic sense. Higher order curves do not make economic sense. Second, there were not enough data to reasonably estimate the parameters for the higher order curve. So, no attempts were made at fitting higher order curves to the data.

Multivariate Regression

If a nonlinear model did not seem reasonable, what about additional variables? Three additional variables were tried: the money supply M1, lagged one year, the metals indices discussed in Chapter IV, and the national unemployment rate. The unemployment rate for the aircraft industry would have been used, but only data after 1964 was available. The results of the regressions, are listed in Table IV. The Durbin Watson statistics, the sign of coefficient for the additional variable, and whether the coefficient was significant at the 95% confidence level are all listed. As expected with the addition of any variable, the Durbin Watson statistic increases. In fact, in two cases the Durbin Watson statistic moved out of the critical region for acceptance of autocorrelation. One case was for Production Avionics Index with GNP_d and the Electrical Machinery Index. The sign of the coefficient in this case makes little economic sense. The negative sign says when the Electrical Index goes up the Production Avionics Index goes down. The other case where the Durbin Watson statistic moved out of the critical region

GOR/SM/76D-9

Table IV

Durbin Watson Statistics

Straight Regressions	GNP_d	GNP_d Plus Metal	GNP_d Plus $M1_{t-1}$	GNP_d Plus Unemployment
Development AF	.25	$.83^{s}_{-}$	$.78^{N}_{+}$	1.23^{s}_{-}
Production AF	.24	$.80^{s}_{-}$	$.71^{N}_{+}$	1.25^{s}_{-}
Development Engine	.65	1.19^{s}_{+}	1.36^{s}_{-}	1.63^{s}_{+}
Production Engine	.42	$.90^{s}_{+}$	1.12^{s}_{-}	1.06^{s}_{+}
Development Avionics	.28	1.41^{s}_{-}	$.85^{N}_{+}$	1.15^{s}_{-}
Production Avionics	.39	1.62^{s}_{-}	1.08^{s}_{+}	1.36^{s}_{-}
D_u	1.39	1.53	1.53	1.53

D_u is from the Durbin Watson tables

+, - indicate the sign of the coefficient for the additional variable

s means the coefficient of the additional variable was statistically significant at the 95% confidence level.

N means the coefficient of the additional variable was not statistically significant.

was for Development Engine Index with GNP_d and unemployment. One might be able to justify, economically, the addition of unemployment if one ignored the signs of the coefficients of the other indices. The signs indicate that unemployment is positively related to the Engine indices and negatively re-

lated to the Airframe and Avionics indices. It would be hard to explain this phenomenon in economic or business terms.

The difference in signs of the coefficients is not hard to explain, statistically. An additional variable will be statistically significant if it relates well to the residuals, the differences between the regression line and the actual values, of the GNP_d only regression. As shown in Figs. 6 to 11, the pattern of the residuals for the Airframe and Avionics indices is opposite that of the Engine indices. The sign of the coefficients reflect these same patterns.

One can even predict whether or not an additional variable will be statistically significant. The pattern of the residuals shows that any additional variable that has a dip or rise in the middle of its data will probably be statistically significant. This is why one must be so careful when doing a regression. Many variables could be added to a regression that are statistically significant, but which are not consistent with the underlying phenomena being modeled. None of the variables tried here are recommended as additional variables in Ordinary Least Squares analysis on the composite indices.

First Order Autoregressive Model

The next technique tried was the assumption of a first order autoregressive model. The Durbin Two Step, the Search,

and the Bayesian techniques were used to try and estimate ρ. These techniques are explained in Chapter III. Computer code for the Search and Bayesian techniques is included in Appendix D. The results of the different techniques are displayed in Table V. As explained in Chapter III, the Search answers were considered the most reliable, the Durbin Two Step answers were used as a check, and the Bayesian answers were used to estimate how much greater than one ρ might be. There was a problem with the Bayesian technique. For some cases the variance did not reach a single minimum, rather the variance reached a local minimum then increased and then decreased to another local minimum. In these cases only the first local minimum reached was displayed. These cases are indicated by a $\sim$ in Table V.

The results of the one variable, Search, Durbin Two Step, and Bayesian techniques are consistent with the raw data. The simple regressions done in the beginning of the chapter, all resulted in very small Durbin Watson statistics. These small Durbin Watson statistics are consistent with a large value of ρ in the first order autoregressive model. The Development Engine index had the largest Durbin Watson statistic for the simple regression. As expected, this index had the smallest estimate of ρ for the one variable Search technique. Also, a close look at Figs. 6 through 11 reveals a slight tendency for the data to be close to the regression line for the small values of the GNP price deflator and any farther away for large values of the GNP price defla-

Table V

Estimates of ρ

	Two Step Durbin GNP_d only	Search GNP_d only	Bayesian GNP_d only	Search GNP_d & Metals	Search GNP_d & General/ Unemployment
Development Airframe	1.04	>.99	1.1∿	>.99-	>.99-
Production Airframe	1.03	>.99	1.1∿	>.99	>.99-
Development Engine	1.01	.95		.71+	.3 +
Production Engine	1.31	>.99	1.1∿	>.99+	>.99+
Development Avionics	.94	>.99	1.1	.35-	>.9 -
Production Avionics	.85	>.99	1.1	.21-	>.9 -

∿ means there was more than one local minimum of the variance. The first minimum reached is listed.

+, - are the signs of the coefficients of the additional variables.

tor. This apparent divergence of the data is consistent with a ρ slightly greater than one in the first order autoregressive model.

The Search technique was also used with an additional variable. Both the general unemployment rate and the metals indices were tried. M1 was not tried because of its high correlation with GNP_d creating dependency problems. See OLS

assumption 4, Cahpter III. The results are displayed in Table V. The signs of the coefficients of the additional variables are also listed in Table V. The signs follow the same pattern as they did earlier in the ordinary linear model regressions.

Partially because of the signs of the coefficients, predictions were only made for two cases, Predictions were made for the Development Engine index for ρ = .95, and for ρ = .71 with the Primary Metals index as an additional variable. The predictions along with the prediction intervals are listed in Table VI. Predictions were not made for the Avionics indices with the addition of the Electrical Machinery index. This was because the coefficient of the Electrical Metals index was negative and made no economic sense. Predictions were not made for the Development Engine index with unemployment added. As explained earlier in the chapter, the pattern of the signs of the coefficients of unemployment in the six indices is difficult to explain economically. The predictions were only made for those cases where Search indicated ρ was less than one. Also predictions were not made for those cases where ρ might have been larger than one because regression theory does not explain how to predict in this case.

As seen in Tables II, III, and VI, the prediction intervals for the first order autoregressive model with Generalized Least Squares showed only a slight improvement over the per cent change regressions. The first order model with

Table VI

Generalized Least Squares Prediction of Developmental Engine Index

	$\hat{\rho}$ = .95 GNP_d only	$\hat{\rho}$ = .71 GNP_d & PM_t
1976	184.02± 4.60	193.19± 4.15
1977	208.00± 6.03	206.89± 5.26
1978	225.66	223.60
1979	245.96	243.35
1980	264.20	259.73
1981	278.50	273.62
1982	297.29	291.89
1983	316.54	310.57
1984	329.80	323.53
1985	339.45±18.22	333.07±11.32

Prediction intervals = 2 $\sigma\rho$

where $\sigma\rho$ is the standard deviation of the prediction

Generalized Least Squares is, however, not recommended over the percent change technique for three reasons. As will be discussed below, the first reason is that there is uncertainty in the assumption of the first autoregressive model. The second reason is that the first order autoregressive model with Generalized Least Squares is much more difficult to use. The third reason is that the first order autoregressive model

would be used for only one of the six indices.

First Differences

The results of the previous section all indicate that ρ is close to one. When ρ is close to one, first differences is recommended. First differences was tried on the six composite indices. However, the results were judged inferior to the per cent change results for two reasons: per cent changes had better residual patterns (see Chapter III), and the first differences regression had a significant constant term. The constant term was statistically significant at the 95% confidence level for all but the Development Engine equation. If the assumption of a first order autoregressive model is correct and if ρ is close to one, then the constant term in the first difference regression should not be significant. The significance of the constant term indicates that the assumption of the first order regressive model with ρ close to one for the model specified with one independent variable, GNP_d, is not valid.

Economic principles reveal why first differences could be an inferior model compared to per cent changes. First differences considers only the change in an index, not the relative change. This is fine when two indices maintain the same relative position to each other, but when one index begins to grow or decline much faster than the other, a simple change may no longer be relevant. In 1958 the GNP price deflator was 66.06 and the Development Airframe was 64.44, but

in 1975 the GNP price deflator was 126.37 and the Development Airframe Index was 176.94. Relative to the GNP price deflator, a change of say 10 in the Development Airframe Index would be less significant in 1975 than in 1958. Per cent changes would recognize this fact, first differences would not.

The Primary Metals Index, the Electrical Machinery Index, the national unemployment rate, and the lagged value of the money supply were all tried as additional variables in the per cent change regressions. None of these variables were recommended as additions to the per cent change model either because they were not statistically significant at the 95% confidence level, or that the signs of their coefficients did not make economic sense.

Conclusions and Recommendations

Simply said, when starting with the composite indices, no technique was found to be superior to per cent change regressions. However, quite different results might be obtained if one started with the subindex data. A quick look at the subindex data shows three sets of patterns. One set consists of the labor indices, one set consists of the material indices, and one set consists of the overhead indices. It is quite probable that one could find additional variables that made economic sense and were also statistically significant for one of the sets of indices.

VII. The Wharton Predictors

Errors in the Predictors

Up to this point the predictors used in the regressions equations were assumed to be without error. The future, however, is never really known without error, and the predictions used in this thesis are no exception. This chapter will discuss the errors in the predictors and how these errors affect the predicted values of the six aeronautical price indices.

The predictors used by ASD and used in this thesis are from the Wharton Mark IV long range econometric forecasting model. The Wharton model consists of about 1000 interrelated equations which attempt to describe the US economy. The model is not a closed mathematical system, rather it needs exogenous inputs. Each quarter a group of businessmen, economists, and other experts meet at Wharton to help determine the subjective inputs of the mathematical model. Besides business data, estimates of national policy variables under the influence of the Federal Reserve, the President, and the Congress, are placed into the model. Wharton does allow for multiple solutions depending upon the user's preference. For example, Wharton provides estimates of the economy for a tight money or easy money scenario. The fact remains, though, that the Wharton estimates depend on both mathematical equations and subjective judgements (Ref 26: Ref 27).

There are models that do not use subjective inputs.

Brush, for instance, proposes a prediction for inflation that uses no (or at least very little) subjective inputs (Ref 28). Brush's model may give good results for two years into the future. This is because all economic policy decisions, especially those of Congress, have a considerable lag time before implementation. Furthermore, implemented policy takes time before its effect is felt. Unfortunately, ASD has a need for estimates up to ten years in the future. The need for long term predictions and a desire for consistency in reporting, caused ASD to choose Wharton's predictors over Brush's.

How good are the Wharton forecasts? That is hard to say. Only one study could be located that objectively evaluated Wharton's predictors (Ref 29). It compared a number of economic forecasting models. Wharton's Mark IV was not included, but rather the older Mark III quarterly version was tested. However, the Wharton model did show the lowest error of all the models tested over a two year period. The errors for the Wharton GNP price deflator for two years in the future were less than 3 percentage points of the 1958 value.

How the Errors Affect the Price Index Predictions

How do the errors in the Wharton predictors affect the errors in the aeronautical price indices? Figs. 12 and 13 will help explain the effect. In this example the Development Engine index will be used. The relationship between

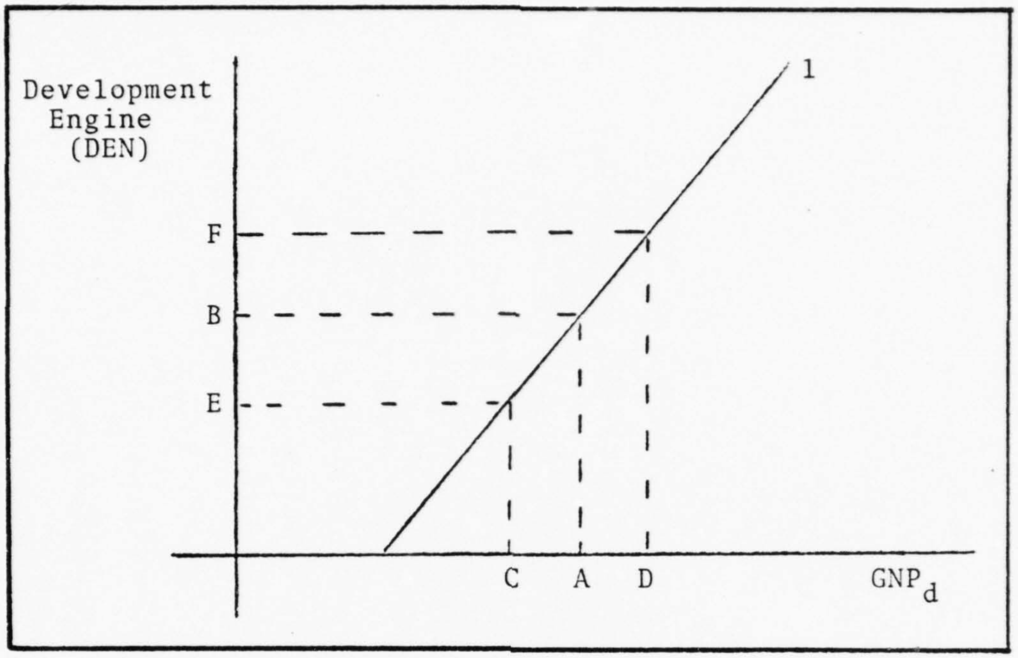


Fig. 12. Partial Affect of Error in GNP_d

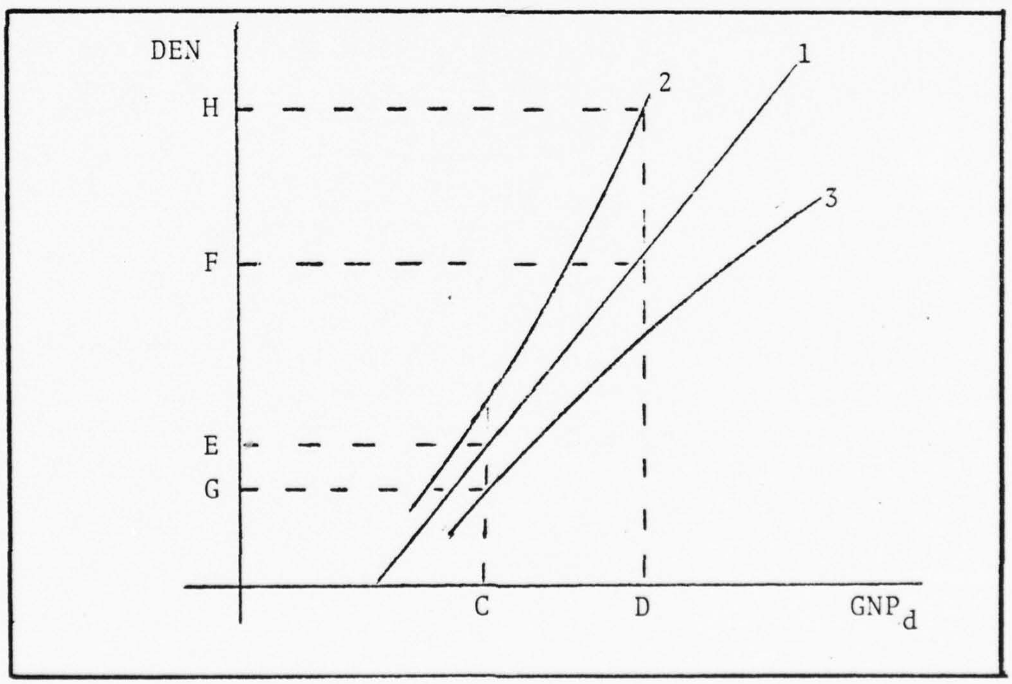


Fig. 13. Total Affect of Error in GNP_d

the Development Engine index (Dev Engine) and the GNP price deflator (GNP_d) is represented by line 1 in Fig. 11. For any GNP_d, e.g. point A, a corresponding estimate of Dev Eggine, point B, can be determined. But what if point A is not known exactly but rather, because of error, is known to be in the interval CD. Then if there were no error in the Dev Engine relationship, the predictors of Dev Engine would be in interval FE. If FE is larger than CD (assuming equal scales) then this is the penalty one pays when he wants to predict a specific index, not just GNP_d.

The relationship between DAF and GNP_d is not known without error. Prediction intervals for their relationship is represented by lines 2 and 3 in Fig. 12. Now, the interval CD results in interval HG, which is better than FE. The phenomena represented here in Figs. 11 and 12, is not quite mathematically exact, but it does represent what really happens.

Adding sample numbers will illustrate the magnitude of the errors. From Table III, the predicted error in 1977 for the Development Engine Index is 6.67. The root mean square error of the Wharton predictor is approximately 3% of the 1958 value of 66.06, which equals 1.98. If one considers the root mean square error as an approximation to the standard deviation of the error, then two times the root mean square error would yield 95% confidence bounds, if the error was normally distributed. Thus the error in GNP_d would be 3.96. The regression line for the indices has an equivalent slope

of close to 1, so the total error for the Development Engine Index in 1977 would be $\pm 1.98 + 6.67 = \pm 8.65$. This is certainly only an estimate, and is not mathematically exact, but it does give an idea of the magnitude of the possible prediction errors.

VIII. Summary, Findings, and Recommended Areas for Future Study

Summary

The thesis began with a description of the annual cost escalation reports prepared by the Aeronautical Systems Division (ASD) of the United States Air Force Systems Command. The reports contain predictions for selected aircraft price indices for up to ten years in the future. The thesis then recalculated much of the historical price index data used by ASD in their reports. Next the price indices were repredicted using the same methodology as used by ASD in their latest report. Prediction intervals were also included. The thesis then examined the question of whether it was better to aggregate the historical data first and then predict, or predict subsets of the historical data and then aggregate the predictions. Finally, advanced statistical techniques were tried in an attempt to improve the predictions.

Findings

The significant findings of this thesis can be separated into four different areas: regression theory, historical data, regression techniques, and prediction intervals.

Regression Theory

The use of the Durbin Watson statistic, D, is incorrectly explained in some textbooks. The significance levels given in the Durbin Watson tables only apply to cases where

$D > D_u$, i.e., when the hypothesis of no autocorrelation is rejected.

The Cochrane-Orcutt iterative method for estimating the parameters of a first order autoregressive model is frequently illustrated without the first row of its transformation. The use of Cochrane-Orcutt without the first row, in cases where autocorrelation has been going on before the first data point, reduces the efficiency of the technique.

The Search technique developed in Chapters III and VI, yields maximum likelihood estimates of the parameters of the first order autoregressive model. The Search technique is recommended by the author as the most accurate technique for estimating the parameters of the first order autoregressive model.

Historical Data

Almost all of the historical data used by ASD in their latest report was examined and recalculated. There were only a few minor differences between the recalculated data and the data listed in the ASD 110C report. The most significant difference was in the calculation of the Airframe Raw Material subindex. The difference was apparently due to a difference in the ASD calculations when the Bureau of Labor Statistics switched from a 1957-1959 base to a 1967 base. ASD has recalculated the Airframe Raw Material Subindex values for use in future reports.

Regression Techniques

Given the independent variables listed in the latest ASD 110C report, the GNP Price Deflator, the Primary Metals index, and the Electrical Machinery index, the following conclusion was made. The aggregate per cent change method is preferred slightly to the per cent change subindex method.

When starting with the composite data, the per cent change regression technique was found to be superior to regression with nonlinear variables, multivariate regression, Generalized Least Squares, and first differences.

Prediction Intervals

Using aggregate per cent change regression, the prediction intervals for the six composite aircraft price indices, given no error in the Wharton predictions, varied widely. However, they averaged approximately 1.8% for the first year, 2.5% for the second year, and 4.7% for the tenth year of the predicted values. Errors in the Wharton predictions would be expected to raise the error in the predictions to a total of 3.5% for the second year.

Recommended Areas for Future Study

The historical data is an important area for future study. The most advanced regression techniques and the most accurate predictions of the independent variables will be useless if the historical data does not accurately represent the true historical price changes. The historical price indices were recalculated in this thesis, but the ASD metho-

dology used to calculate the indices was assumed correct. The methodology should be carefully examined.

The composite indices were used as a starting point when investigating other possible regression techniques. The same investigation could be done starting with the subindex data.

Finally, the variance of the predictions was calculated, but the probability distribution of the predictions was unknown. The normal distribution was assumed in order to calculate prediction intervals. Monte Carlo techniques could be used to construct sample probability distributions leading to perhaps more accurate prediction intervals.

Bibliography

1. Lentzsch, Craig and William D. Bandt. Cost Research Report No. 110A, Historical and Forecasted Aeronautical Cost Indices. Directorate of Cost Analysis, Comptroller, Aeronautical Systems Division, Wright-Patterson Air Force Base, Ohio 45433. May 1973, Revised June 1974.

2. Jackson, Bobby and Craig Lentzsch. Cost Research Report No. 110B, Aeronautical Economic Escalation Indices. Ibid. July 1975

3. -----. ASD Cost Escalation Report No. 110C, Historical and Forecasted Aeronautical Cost Indices. Ibid. April 1976.

4. Campbell, H. G. Aerospace Price Indexes. A report prepared for the United States Air Force Project Rand (R-568-Pr). Rand Corporation, Santa Monica, California. December 1970.

5. Theil, Henri. Principles of Econometrics. John Wiley and Sons, Inc. New York, 1971.

6. U.S. Department of Commerce, National Bureau of Standards (NBS). OMNITAB II User's Reference Manual. Technical note 552. U.S. Government Printing Office, Washington, D.C. 1971.

7. Neter, John and William Wasserman. Applied Linear Statistic Models. Richard D. Irwin, Inc. Homewood, Illinois. 1974.

8. Neary, Peter. "The Relative Efficiency of Regression using Original Data or First Differences: the case of Autocorrelated Disturbances." Economic and Social Review, 5:47-58 (October 1973).

9. Cochrane, D. and G. H. Orcutt. "Application of Least Squares to Relationships Containing Autocorrelated Error Terms." Journal of American Statistical Association, 44:32-61 (March 1949).

10. Kadiyala, Koteswara Rao. "A Transformation Used to Circumvent the Problem of Autocorrelation." Econometrica, 36:92-96 (January 1968).

11. Potluri, Rao and Zvi Griliches. "Small Sample Properties of Several Two-Stage Regression Methods in the Context of Autocorrelated Errors." Journal of the American Statistical Association, 64:253-272 (1969).

12. Durbin, J. "Estimation of Parameters in Time-series Regression Models." Royal Statistical Society Journal Series B, 22:139-153 (January 1960).

13. Dhrymes, P. J. "On Treatment of Certain Recurrent Non-Linearities in Regression Analysis." Southern Economic Journal, 33:187-196 (1966).

14. Hildreth, Clifford, "Asymptotic Disturbances of Max Likelihood Estimates in Linear Models with Autoregressive Disturbances." Rand Memorandum RM-5059-PR. Rand Corporation, Santa Monica, California, 1966.

15. Goldberger, Arthur S. "Best Linear Unbiased Prediction in the Generalized Linear Regressive Model." Journal of the American Statistical Association, 57:369-375 (1962).

16. Zellner, Arnold. An Introduction to Bayesian Inferences in Econometrics. John Wiley and Sons. New York, 1971.

17. -----. Economic Indices for Avionics Equipment. Report No. NADC-SD-7014. Naval Air Development Center, Johnsville. Warminster, Pennsylvania. April 3, 1970.

18. -----. Economic Indices for Avionics Equipment Updated to 1973. Addendum No. 4 to Report No. NADC-SD-7014. Naval Air Development Center, Johnsville. Ibid. 15 May 1974.

19. -----. Wholesale Price Index and Indexes. Bureau of Labor Statistics United States Department of Labor. Washington, D.C.

20. -----. "Selected Wholesale Price Indices with Base Year 1967 = 100." Special computer listing sent from Branch of Wholesale Price Indexes. Bureau of Labor Statistics, U.S. Department of Labor. Washington, D.C. August 1976.

21. -----. Employment and Earnings. Bureau of Labor Statistics. U.S. Department of Labor. Washington, D.C.

22. Benoy, David. Cost Researcher in the Directorate of Cost Analysis Aeronautical Systems Division (ASD). Interview by Author. Headquarters ASD, Wright-Patterson Air Force Base, Ohio. November 24, 1976.

23. Spencer, Richard Allan. A Theoretical Study of Index Number Construction for DoD Use. Thesis. Air Force Institute of Technology. Wright-Patterson Air Force Base, Ohio 45433. June 1971.

24. Whitmore, Robert. Analyst Unemployment/Employment Division Bureau of Labor. Washington, D.C. Phone 523-1944. Phone interview by Author. November 5, 1976.

25. -----. Economic Report of the President. Transmitted to Congress January 1976. U.S. Government Printing Office. Washington, D.C.

26. McLaughlin, Patricia. "The Age of Klein." Pennsylvania Gazette. pp 30-34 (February 1975).

27. McLaughlin, Patricia. "Wharton Predicts." Pennsylvania Gazette. pp 27-30 (March 1975).

28. Brush, John S. A Disequilibrium Adjustment Inflation Forecasting Model. USAFA-TR-75-4. Department of Economics, Geography, and Management. United States Air Force Academy, Colorado Springs, Colorado. July 1975.

29. Christ, C. F. "Judging the Performance of Econometric Models of the US Economy." International Economic Review, 16:57-74 (February 1975).

30. Mendenhall, William and Richard L. Scheaffer. Mathematical Statistics with Applications. Duxbury Press, North Scituate, Massachusetts, 1973.

Appendix A.

Historical Data

Airframe

Subindices

CY YR	Raw Material	AC Parts & Equip.	7.3% Overhead	Composite Raw Mat.	Manufac- turing Labor	Engineering Labor
1958	100.31	72.84	53.04	76.403	71.92	77.52
1959	96.58	76.12	56.91	77.100	75.64	79.02
1960	99.67	78.80	61.07	80.499	77.65	82.55
1961	99.77	80.60	65.52	82.613	79.66	84.23
1962	98.96	83.58	70.31	84.756	82.23	83.72
1963	95.49	86.27	75.44	85.897	84.53	86.41
1964	96.61	88.96	80.95	89.016	85.96	91.44
1965	96.90	91.94	86.86	92.019	90.26	94.46
1966	97.89	95.82	93.20	95.666	95.70	95.64
1967	100.00	100.00	100.00	100.000	100.00	100.00
1968	101.70	105.37	107.30	104.626	104.30	107.21
1969	107.92	112.24	115.13	111.596	111.75	115.10
1970	114.28	119.10	123.54	118.837	119.48	125.34
1971	113.60	124.48	132.56	123.158	126.65	130.03
1972	113.32	132.54	142.23	128.486	138.97	137.08
1973	116.32	139.70	152.61	135.189	146.99	144.96
1974	148.53	149.55	163.75	154.467	159.60	157.05
1975	165.92	164.78	175.71	169.335	177.65	171.81

Airframe (Cont'd)

	Subindex	Aggregate Indices	
	7.8% Overhead	Development Airframe	Production Airframe
1958	50.87	64.44	64.59
1959	54.83	67.22	67.35
1960	59.11	70.84	70.86
1961	63.72	73.98	74.00
1962	68.69	76.90	77.27
1963	74.05	80.44	80.56
1964	79.83	84.93	84.60
1965	86.05	89.58	89.25
1966	92.76	94.34	94.42
1967	100.00	100.00	100.00
1968	107.80	106.56	106.09
1969	116.21	114.45	113.83
1970	125.27	123.22	122.15
1971	135.04	130.49	129.41
1972	145.58	139.54	138.50
1973	156.93	148.78	147.54
1974	169.17	162.40	161.98
1975	182.37	176.94	176.77

Engine

	Subindices				Aggregate	
CY YR	Raw Material	Engine Labor	Purchased Parts	Engineering Labor	Development Engine	Production Engine
1958	112.62	73.39	89.082	77.52	85.46	88.39
1959	108.44	77.19	89.690	79.02	86.31	89.07
1960	106.58	79.82	90.524	82.55	88.02	90.04
1961	107.10	82.16	92.136	84.23	89.64	91.67
1962	105.80	85.09	93.374	83.72	90.25	92.86
1963	100.89	87.43	92.814	86.41	90.74	92.47
1964	99.02	90.35	93.818	91.44	93.07	93.67
1965	99.54	92.69	95.430	94.66	95.15	95.35
1966	98.31	97.08	97.572	95.64	96.92	97.48
1967	100.00	100.00	100.000	100.00	100.00	100.00
1968	102.88	106.72	105.184	107.21	105.84	105.28
1969	107.51	113.16	110.900	115.10	112.28	111.10
1970	118.82	119.88	119.456	125.34	121.45	119.70
1971	122.17	128.07	125.710	130.03	127.13	125.92
1972	123.36	139.18	132.852	137.08	134.16	133.11
1973	135.27	147.95	142.878	144.96	143.48	143.03
1974	165.38	158.77	161.414	157.05	159.97	161.19
1975	200.18	176.32	185.864	171.81	181.26	185.15

Avionics

		Subindices				Aggregate	
CY YR	Raw Material	Development Labor	Development Overhead	Production Labor	Production Overhead	Development Avionics	Production Avionics
1958	105.4	68.82	61.12	71.13	63.17	68.243	74.004
1959	104.1	71.96	64.76	74.28	66.84	71.214	76.524
1960	102.8	74.71	68.12	76.64	69.88	73.894	78.492
1961	100.9	78.04	72.10	80.05	73.96	77.059	81.175
1962	100.6	81.18	76.00	82.41	77.15	80.273	83.418
1963	100.0	84.51	80.17	85.56	81.16	83.672	86.248
1964	97.0	87.65	84.25	88.19	84.77	86.715	88.242
1965	96.0	90.39	88.04	91.08	88.71	89.658	90.879
1966	98.4	93.14	91.92	93.70	92.47	92.995	94.025
1967	100.0	100.00	100.00	100.00	100.00	100.000	100.000
1968	101.3	106.86	108.28	106.30	107.71	107.085	106.005
1969	103.4	113.33	116.36	112.60	115.61	114.003	112.265
1970	105.2	120.00	124.85	119.42	124.24	121.187	118.986
1971	108.2	126.67	133.54	125.46	132.26	128.601	125.408
1972	109.6	132.94	142.00	131.76	140.75	135.589	131.823
1973	110.3	140.00	151.52	140.42	151.98	143.366	140.176
1974	118.1	148.69	163.07	148.92	163.32	153.540	149.956
1975	127.8	162.12	180.16	163.04	181.18	168.610	165.062

Appendix B.
Predictor Variables

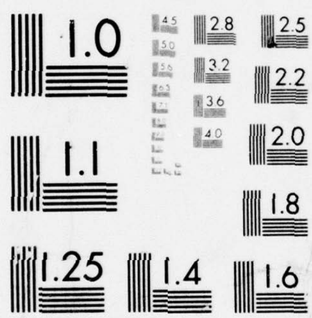
CY YR	GNP Price Deflator	Primary Metals	Electrical Metals	(Billions) Money Supply M1	General Unemployment	Aircraft Industry Unemployment	7.8% Overhead	7.3% Overhead
1957				135.9				
1958	66.06	87.26	113.25	141.1	6.8%			
1959	67.52	88.57	114.27	143.4	5.5%			
1960	68.67	91.54	112.00	144.2	5.5%			
1961	69.28	90.05	111.10	148.7	6.7%			
1962	70.55	92.24	107.25	150.9	5.5%			
1963	71.59	91.71	104.53	156.5	5.7%			
1964	72.71	92.85	102.15	163.7	5.2%	4.5%		
1965	74.32	93.46	99.21	171.3	4.5%	3.7%		
1966	76.76	95.20	97.73	175.4	3.8%	2.1%		
1967	79.02	100.00	100.00	186.9	3.8%	2.0%		
1968	82.57	105.85	102.72	201.7	3.6%	1.9%		
1969	86.72	109.60	98.64	208.7	3.5%	3.3%		
1970	91.36	113.88	103.17	221.4	4.9%	7.6%		
1971	96.02	120.34	105.21	235.3	5.9%	10.8%		
1972	100.00	127.58	105.10	255.8	5.6%	7.0%		
1973	105.92	126.61	106.23	271.5	4.9%	3.1%		
1974	116.20	144.34	113.25	284.4	5.6%	3.0%		
1975	126.37	167.97	124.58	296.4	8.5%	6.8%		
1976	133.86	178.59	135.90				196.59	188.54
1977	142.08	192.03	147.23				211.93	202.30
1978	152.48	206.36	158.55				228.46	217.07
1979	164.45	220.72	169.88				246.28	232.91
1980	175.20	235.74	181.20				265.49	249.92
1981	183.62	248.38	192.53				286.20	268.16
1982	194.70	264.89	203.85				308.52	287.74
1983	206.06	281.58	215.18				332.58	308.74
1984	213.87	293.47	226.50				358.52	331.28
1985	219.55	302.53	237.83				386.49	355.46

AD-A036 463 AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OHIO SCH--ETC F/G 5/3
AN INVESTIGATION OF THE FORECASTING OF AERONAUTICAL PRICE INDIC--ETC(U)
DEC 76 T P LENNERTZ

UNCLASSIFIED GOR/SM/76D-9 NL

2 OF 2
AD
A036463

END
DATE
FILMED
3-77



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Appendix C

Calculation of Prediction Intervals for Per Cent Change Regressions

This appendix shows the theory used to estimate the prediction intervals that are contained in Chapter V. The predictions are for the aeronautical indices that were derived by two methods: the subindex per cent change method, and the aggregate per cent change method. Calculation of prediction intervals for the subindex method required six steps, while calculation of prediction intervals for the aggregate method required only four steps.

Prediction Intervals for the Subindex Method

Step 1 Check Assumptions of Ordinary Least Squares

The assumptions of Ordinary Least Squares, discussed in Chapter III, were checked for each regression equation. This ensured that the errors within each regression were independent and were normally distributed.

Step 2 Calculate Variances for the Per Cent Change Prediction

The Wharton perdictors were substituted in to the regression equations to yield predictions for the per cent change of the subindices. The variances for these per cent change predictions were easily calculated using the following formula:

$$\mathrm{Var}(\Delta\%_Y) = \sigma^2(1 + \alpha'_{kx1}\,(X'_{nxk}\,X_{nxk})^{-1}\alpha_{kx1}) \qquad (64)$$

where Var($\Delta\%_Y$) is the variance of the per cent change for year y. σ^2 is estimated by S^2, and α_{kx1} is the vector of predictors for year y (Ref 5:135). The results for the engineering labor subindex were:

$$\% \text{ Engineering Labor} = 1.582 + .8286\ \Delta\%GNP_d \tag{65}$$

$$\Delta\%_{76} = 6.4937 \tag{66}$$

$$Var(\Delta\%_{76}) = 4.25339 \tag{67}$$

$$\Delta\%_{77} = 6.6707 \tag{68}$$

$$Var(\Delta\%_{77}) = 4.287 \tag{69}$$

Step 3 Calculate Variances of the Subindex Predictions

The predictions are easily calculated using Eqs (50) and (51). The variances are more difficult to calculate. The following is an example of the calculations for the engineering labor subindex along with a brief explanation:

$$1976 = \left(\frac{\Delta\%_{76}}{100} + 1\right)(1975) = \left(\frac{6.4937}{100} + 1\right)(171.81) = 182.97 \tag{70}$$

$$Var(1976) = \left(\frac{1975}{100}\right)^2 Var(\Delta\%_{76}) = \left(\frac{171.81}{100}\right)^2 (4.25339) = 12.558 \tag{71}$$

$$1977 = \left(\frac{\Delta\%_{77}}{100^2} + 1\right)(1976) = \left(\frac{6.6707}{100^2} + 1\right)(182.97) = 195.17 \tag{72}$$

$$Var(1977) = Var\left[\left(\frac{\Delta\%_{77}}{100} + 1\right)(1976)\right] = \frac{1}{100^2} Var[(\Delta\%_{77})(1976)] +$$

$$Var\ 1976 = \frac{1}{100^2}\left[(1976)^2 Var(\Delta\%_{77}) + (\Delta\%_{76})^2 Var\ (1976) +\right.$$

$$\left. Var(1976)\ Var(\Delta\%_{76})\right] + Var(1976) = \frac{1}{100^2}\left[(182.97)^2(4.287) +\right.$$

$$\left.(6.6707)^2(12.555) + 112.555)(4.287)\right] + 12.55 = \underline{26.968} \tag{73}$$

where 1976 means the 1976 prediction of the subindex. Eq (73) uses the assumption that the 1975 index is known without error and that the predictors are independent of each other. Eq (73) is based on the following relationship:

if X and Y are independent random variables then

$$\mathrm{Var}(XY) = E(X)^2\ \mathrm{Var}(X) + E(Y)^2\ \mathrm{Var}(Y) + \mathrm{Var}(X)\ \mathrm{Var}(Y) \tag{74}$$

where E(X) is the expected value of X.

Step 4 Estimate the Covariance Between the Subindices

The covariances are needed for Step 5. They are estimated using the sample correlation matrix of the residuals of the subindex regressions adjusted each year for the variances of the subindex predictions. The following is an example:

sample correlation between subindex 5 (Avionics Development Labor) and subindex 9 (Avionics Raw Material)

$$= \hat{\rho}_{59} = -.2444 \tag{75}$$

$$\text{Variance for subindex 5, 1976} = \sigma_5{}^2 = 4.027 \tag{76}$$

$$\text{Variance for subindex 9, 1976} = \sigma_9{}^2 = 7.809 \tag{77}$$

$$\text{Covariance between 5 and 9} = \sigma_{59} = -.2444(\sqrt{4.027})(\sqrt{7.809}) = -1.378 \tag{78}$$

Step 5 Determine Variances of Composite Indices

The composite indices are simply a linear combination of subindices as shown in Eqs (19) through (26). The variances of a linear combination of interdependent random variables is calculated as follows:

if Y = aX + bX

then $\mathrm{Var}(Y) = a^2\sigma^2_{xx} + 2ab\sigma_{xz} = b^2\sigma^2_{zz}$ (79)

where X, Y, Z are random variables. a, b are constants and σ^2_x is the variance of X, and σ_{xz} is the covariance between X and Z.

So for the Avionics Development Index

$$\mathrm{Var}(1976) = (.9693)^2(4.027) + (.1)^2(7.889) + 2(.1)(.9693)(-1.378) = 3.595 \quad (80)$$

<u>Step 6 Calculate Prediction Intervals</u>

Step 1 started with normally distributed random errors. The predictions in Step 2 were still normally distributed, but in Step 3 the predictions were no longer normally distributed. In Step 4 only estimates of the covariances were available. So the variances in Step 5 are only estimates and they are for unknown probability distributions. Still it is felt that two times the standard deviation would be close to a 95% prediction interval (Ref 30:8). Thus for the Avionics Development Index a 95% prediction interval for 1976 is

$$= 180.27 \pm 2\sqrt{3.595}$$

$$= 180.27 \pm 3.7 \quad (81)$$

Prediction Intervals for the Aggregate Method

<u>Step 1 Check Assumptions of Ordinary Least Squares</u>

This is exactly the same as for the subindex method, except now the six aggregate regressions are checked.

Step 2 Calculate Variances for Index Predictions

As in Step 2 of the subindex method, the Wharton predictors were substituted in the regression equations to yield the prediction for the indices. The variances were again easily calculated. For example, for the Development Airframe Index:

$$\Delta\%\ \text{Development Airframe (DAF)} = 3.959 + .5544\ \Delta\%\ GNP_d \tag{82}$$

$$\Delta\%_{76} = 7.2455$$

$$Var(\Delta\%_{76}) = .42362$$

$$\Delta\%_{77} = 7.3640$$

$$Var(\Delta\%_{77}) = .42701$$

Step 3 Calculate Variances of Index Predictions

The same mathematical theory that was used in Step 3 of the subindex method is used in this step. Examples for the Development Airframe Index are:

$$1976 = 176.9454\ (\frac{7.2455}{100} + 1) = 189.765 \tag{83}$$

$$Var(1976) = (\frac{176.9454}{100})^2\ (.42362) = 1.326 \tag{84}$$

$$1977 = 189.765\ (\frac{7.364}{100} + 1) = 203.739 \tag{85}$$

$$Var(1977) = \frac{1}{100^2}\ [(189.765)^2(.42701) + (7.364)^2(1.326) + (1.326)(.42701)] + 1.326 = \underline{2.87} \tag{86}$$

Step 4 Calculate Prediction Intervals

Since the data was already aggregated, covariances are not needed. Through Step 2 the variances were normally dis-

tributed. However, the variances from Step 3 are not normally distributed, but it is still felt that two times the standard deviation will yield good estimates for 95% prediction intervals. For example, the prediction interval for the 1976 Development Airframe index is:

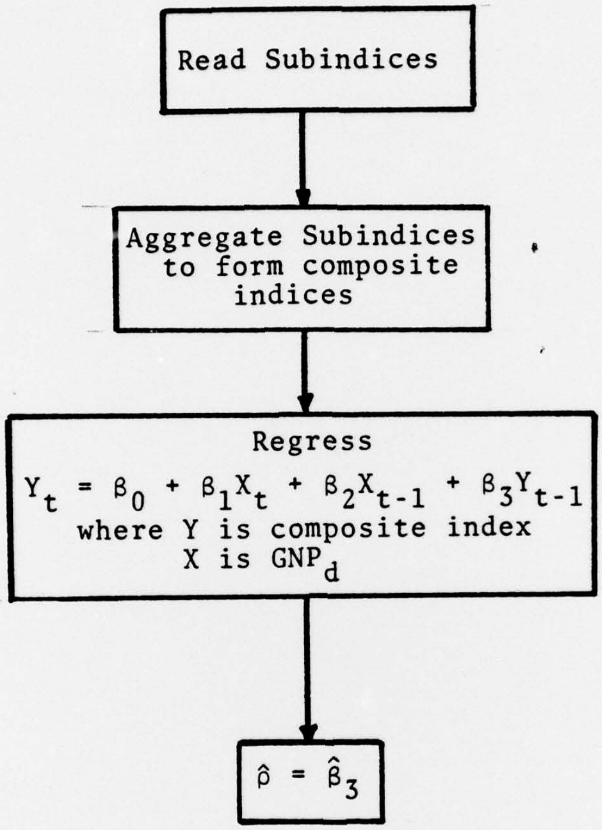
$$189.765 \pm 2\sqrt{1.326}$$

$$189.765 \pm 2.3 \qquad (87)$$

Appendix D

Flow Charts and Computer Codes

Durbin Two Step Flow Chart



Durbin Two Step Computer Code (Ref 6)

Read in subindices

```
OMNITAB

DIMENSION 28X42
SET ENGINEERING LABOR IN 1
77.52 79.02 82.55 84.23 83.72 86.41 91.44 94.46 95.64 100.
107.21 115.1 125.34 130.03 137.08 144.96 157.05 171.81
SET MANUFACT LABOR IN COL 2
71.92 75.64 77.65 79.66 82.23 84.53 85.96 90.26 95.70 100.
104.30 111.75 119.48 126.65 138.97 146.99 159.60 177.65
SET AC PARTS AND EQUIP IN COL 3
72.84 76.12 78.80 80.60 83.58 86.27 88.96 91.94 95.82 100.
105.37 112.24 119.10 124.48 132.54 139.70 149.55 164.78
SET AC ENGINE LABOR IN COL 4
73.39 77.19 79.82 82.16 85.09 87.43 90.35 92.69 97.08 100.
106.72 113.16 119.88 128.07 139.18 147.95 158.77 176.32
SET AVIONIC DEV LABOR IN COL 5
68.82 71.96 74.71 78.04 81.18 84.51 87.65 90.39 93.14 100.
106.86 113.33 120.00 126.67 132.94 140.00 148.69 162.12
SET AVIONIC PROD LABOR IN COL 6
71.13 74.28 76.64 80.05 82.41 85.56 88.19 91.08 93.70 100.
106.30 112.60 119.42 125.46 131.76 140.42 148.92 163.04
SET AF RAW MAT IN COL 7
100.31 96.58 99.67 99.77 98.96 95.49 96.61 96.90 97.89 100.
101.70 107.92 114.28 113.60 113.32 116.32 148.53 165.92
SET ENGINE RAW IN COL 8
112.62 108.44 106.58 107.10 105.80 100.89 99.02 99.54 98.31
100. 102.88 107.51 118.82 122.17 123.36 135.27 165.38 200.18
SET AVIONIC RAW MAT IN COL 9
105.4 104.1 102.8 100.9 100.6 100. 97.0 96. 98.4 100.
101.3 103.4 105.2 108.2 109.6 110.3 118.1 127.8
SET AVIONIC DEV LABOR IN COL 5
68.82 71.96 74.71 78.04 81.18 84.51 87.65 90.39 93.14 100.
106.86 113.33 120.00 126.67 132.94 140.00 148.69 162.12
SET AVIONIC PROD LABOR IN COL 6
71.13 74.28 76.64 80.05 82.41 85.56 88.19 91.08 93.70 100.
106.30 112.60 119.42 125.46 131.76 140.42 148.92 163.04
SET AF RAW MAT IN COL 7
100.31 96.58 99.67 99.77 98.96 95.49 96.61 96.90 97.89 100.
101.70 107.92 114.28 113.60 113.32 116.32 148.53 165.92
SET ENGINE RAW IN COL 8
112.62 108.44 106.58 107.10 105.80 100.89 99.02 99.54 98.31
100. 102.88 107.51 118.82 122.17 123.36 135.27 165.38 200.18
SET AVIONIC RAW MAT IN COL 9
105.4 104.1 102.8 100.9 100.6 100. 97.0 96. 98.4 100.
101.3 103.4 105.2 108.2 109.6 110.3 118.1 127.8
```

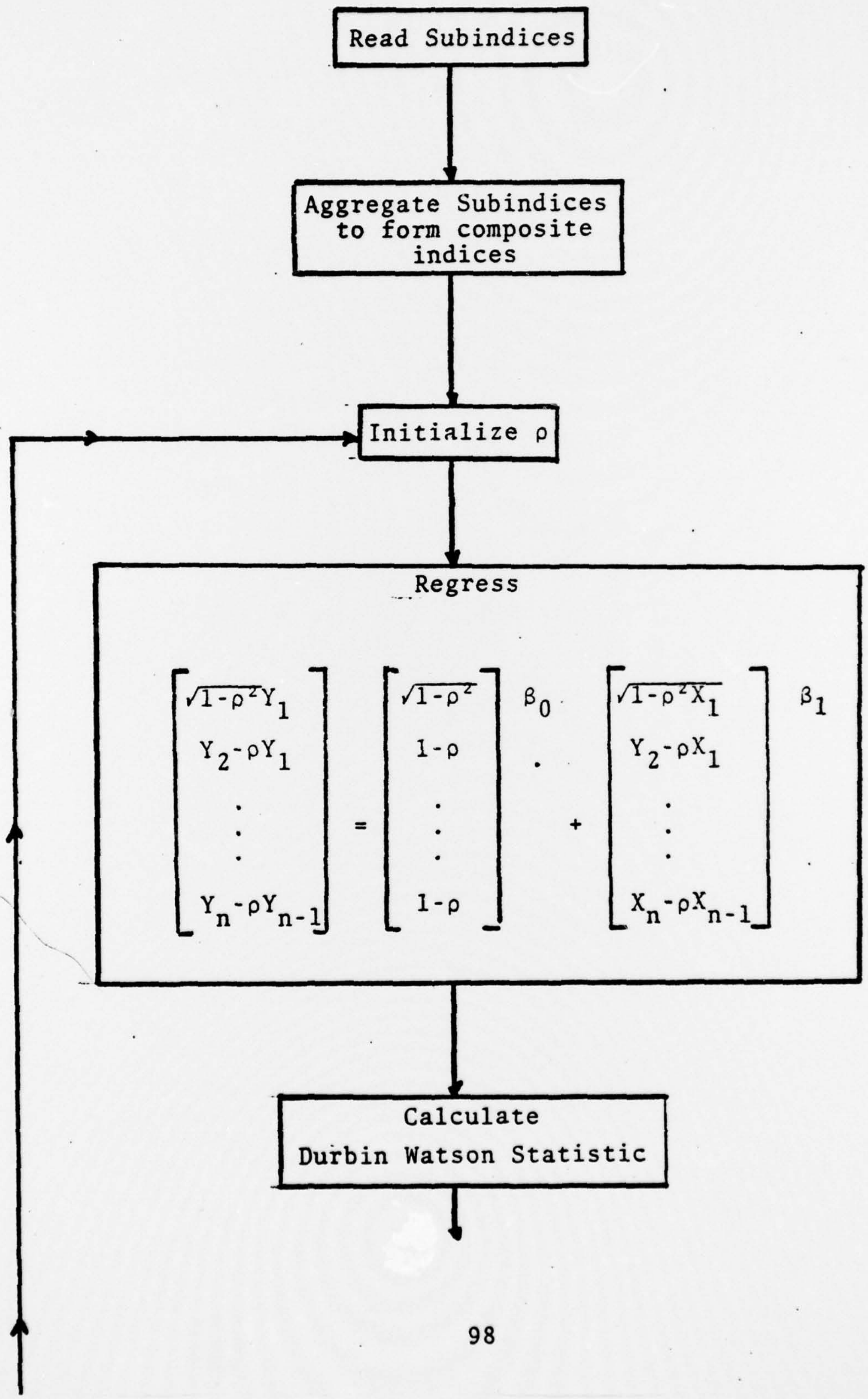
```
SET SEVEN PT THREE OVERHEAD IN COL 10
53.04 56.91 61.07 65.52 70.31 75.44 80.95 86.86 93.20 100.
107.30 115.13 123.54 132.56 142.23 152.61 163.75 175.71
SET SEVEN PT EIGHT OVERHEAD IN 11
50.87 54.83 59.11 63.72 68.69 74.05 79.83 86.05 92.76 100.
107.8 116.21 125.27 135.04 145.58 156.93 169.17 182.37
SET AVIONICS PROD OH IN 13
63.17 66.84 69.88 73.96 77.15 81.16 84.77 88.71 92.47 100.
107.71 115.61 124.24 132.26 140.75 151.98 163.32 181.18
SET AVIONICS DEV OH IN 12
61.12 64.76 68.12 72.1 76. 80.17 84.25 88.04 91.92 100.
108.28 116.36 124.85 133.54 142. 151.52 163.07 180.16
```

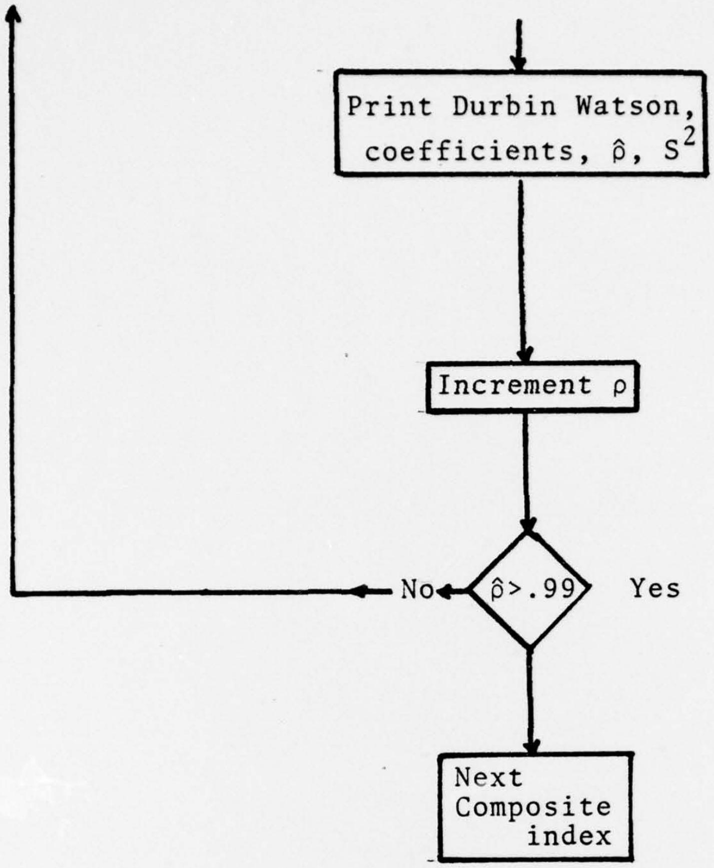
Aggregate data and Regress

```
MULT 7 BY .4 PUT IN 14
MULT 3 BY .225 BULT BY 1. ADD 14 PUT IN 14
MULT 10 BY .375 1. 14 14
MULT 2 BY .1425 PUT IN 19
MULT 1 BY .2125 1. 19 19
MULT 14 BY .1924 1. 19 19
MULT 11 BY .4526 1. 19 19
MULT 2 BY .2010 PUT IN 20
MULT 1 BY .0650 1. 20 20
MULT 14 BY .3040 1. 20 20
MULT 11 BY .4300 1. 20 20
MULT 8 BY .4 PUT IN 15
MULT 4 BY .6 BY 1. 15 15
MULT 4 BY .2701 PUT IN 21
MULT 1 BY .3415 1. 21 21
MULT 8 BY .1942 1. 21 21
MULT 15 BY .1942 1. 21 21
MULT 4 BY .4184 PUT IN 22
MULT 1 BY .0416 1. 22 22
MULT 8 BY .2700 1. 22 22
MULT 15 BY .2700 1. 22 22
MULT 5 BY .35 PUT IN 23
MULT 9 BY .1 1. 23 23
MULT 12 BY .55 1. 23 23
MULT 6 BY .3 PUT IN 24
MULT 9 BY .2 1. 24 24
MULT 13 BY .5 1. 24 24
MOVE 1,19 18X6 TO 1,1
MOVE 2,1 17X9 TO 1,10
SET GNP DEFLATOR IN COL 30
66.06 67.52 68.67 69.28 70.55 71.59 72.71 74.32 76.76 79.02
82.57 86.72 91.36 96.02 100. 105.92 116.2 126.37
DEFINE 1.0 IN COL 28
ERASE 7***27
```

```
MOVE 2,1 17X6 TO 1,7
MOVE 2,30 17X1 TO 1,31
SPACE 3
MPRINT 1,1 18X12
MPRINT 1,30 18X2
RESET 17
BEGIN
1FIT Y IN 10 WTS 1.0 4 VAR IN 28,31,30,4 C 33
2SPACE 3
3MPRINT 1,33 18X1
4INCREMENT INSTRUCT 1 BY 1,0.,0,0,0,0,1,0
FINISH
PERFORM 1 TO 4 6 TIMES
STOP
```

Search Flow Chart





Print Durbin Watson, coefficients, $\hat{\rho}$, S^2
Increment ρ
No
$\hat{\rho}>.99$
Yes
Next Composite index

Search Computer Code (Ref 6)

Read in subindices

```
OMNITAB

DIMENSION 28X42
SET ENGINEERING LABOR IN 1
77.52 79.02 82.55 84.23 83.72 86.41 91.44 94.46 95.64 100.
107.21 115.1 125.34 130.03 137.08 144.96 157.05 171.81
SET MANUFACT LABOR IN COL 2
71.92 75.64 77.65 79.66 82.23 84.53 85.96 90.26 95.70 100.
104.30 111.75 119.48 126.65 138.97 146.99 159.60 177.65
SET AC PARTS AND EQUIP IN COL 3
72.84 76.12 78.80 80.60 83.58 86.27 88.96 91.94 95.82 100.
105.37 112.24 119.10 124.48 132.54 139.70 149.55 164.78
SET AC ENGINE LABOR IN COL 4
73.39 77.19 79.82 82.16 85.09 87.43 90.35 92.69 97.08 100.
106.72 113.16 119.88 128.07 139.18 147.95 158.77 176.32
SET AVIONIC DEV LABOR IN COL 5
68.82 71.96 74.71 78.04 81.18 84.51 87.65 90.39 93.14 100.
106.86 113.33 120.00 126.67 132.94 140.00 148.69 162.12
SET AVIONIC PROD LABOR IN COL 6
71.13 74.28 76.64 80.05 82.41 85.56 88.19 91.08 93.70 100.
106.30 112.60 119.42 125.46 131.76 140.42 148.92 163.04
SET AF RAW MAT IN COL 7
100.31 96.58 99.67 99.77 98.96 95.49 96.61 96.90 97.89 100.
101.70 107.92 114.28 113.60 113.32 116.32 148.53 165.92
SET ENGINE RAW IN COL 8
112.62 108.44 106.58 107.10 105.80 100.89 99.02 99.54 98.31
100. 102.88 107.51 118.82 122.17 123.36 135.27 165.38 200.18
SET AVIONIC RAW MAT IN COL 9
105.4 104.1 102.8 100.9 100.6 100. 97.0 96. 98.4 100.
101.3 103.4 105.2 108.2 109.6 110.3 118.1 127.8
SET PRIMARY METAL DEFLATOR IN COL 41
87.26 88.57 91.54 90.05 92.24 91.71 92.85 93.46 95.20 100.
105.85 109.6 113.88 120.34 127.58 126.61 144.34 167.97
178.59 192.03 206.36 220.72 235.74 248.38 264.89 281.58
293.47 302.53
SET ELECT METAL DEFLATOR IN COL 42
113.25 114.27 112. 111.1 107.25 104.53 102.15 99.21 97.73 100.
 102.72 98.64 103.17 105.21 105.1 106.25 113.25 124.58
135.90 147.23 158.55 169.88 181.20 192.53 203.85
215.18 226.5 237.83
SET SEVEN PT THREE OVERHEAD IN COL 10
53.04 56.91 61.07 65.52 70.31 75.44 80.95 86.86 93.20 100.
107.30 115.13 123.54 132.56 142.23 152.61 163.75 175.71
SET SEVEN PT EIGHT OVERHEAD IN 11
50.87 54.83 59.11 63.72 68.69 74.05 79.83 86.05 92.76 100.
107.8 116.21 125.27 135.04 145.58 156.93 169.17 182.37
```

```
SFT AVIONICS PROD OH IN 13
63.17 66.84 69.88 73.96 77.15 81.16 84.77 88.71 92.47 100.
107.71 115.61 124.24 132.26 140.75 151.98 163.32 181.18
SFT AVIONICS DEV OH IN 12
61.12 64.76 68.12 72.10 76. 80.17 84.25 88.04 91.92 100.
108.28 116.36 124.85 133.54 142. 151.52 163.07 180.16
```

Aggregate data

```
MULT 7 BY .4 PUT IN 14
MULT 3 BY .225 BULT BY 1. ADD 14 PUT IN 14
MULT 10 BY .375 1. 14 14
MULT 2 BY .1425 PUT IN 19
MULT 1 BY .2125 1. 19 19
MULT 14 BY .1924 1. 19 19
MULT 11 BY .4526 1. 19 19
MULT 2 BY .2010 PUT IN 20
MULT 1 BY .0650 1. 20 20
MULT 14 BY .3040 1. 20 20
MULT 11 BY .4300 1. 20 20
MULT 8 BY .4 PUT IN 15
MULT 4 BY .6 BY 1. 15 15
MULT 4 BY .2701 PUT IN 21
MULT 1 BY .3415 1. 21 21
MULT 8 BY .1942 1. 21 21
MULT 15 BY .1942 1. 21 21
MULT 4 BY .4184 PUT IN 22
MULT 1 BY .0416 1. 22 22
MULT 8 BY .2700 1. 22 22
MULT 15 BY .2700 1. 22 22
MULT 5 BY .35 PUT IN 23
MULT 9 BY .1 1. 23 23
MULT 12 BY .55 1. 23 23
MULT 6 BY .3 PUT IN 24
MULT 9 BY .2 1. 24 24
MULT 13 BY .5 1. 24 24
MOVE 1,19 18X6 TO 1,1
ERASE 7***24
SET UNEMPLOYMENT IN COL 18
6.8 5.5 5.5 6.7 5.5 5.7 5.2 4.5 3.8 3.8
3.6 3.5 4.9 5.9 5.6 4.9 5.6 8.5
SET LAGGED M ONE IN COL 17
135.9 141.1 143.4 144.2 148.7 150.9
156.5 163.7 171.3 175.4 186.9 201.7 208.7
221.4 235.3 255.8 271.5 284.4
```

```
RESET 28
DEFINE 1.0 IN COL28
SET GNP DEFLATOR IN COL 30
66.06 67.52 68.67 69.28 70.55 71.59 72.71 74.32 76.76 79.02
82.57 86.72 91.36 96.02 100. 105.92 116.2 126.37
133.86 142.08 152.48 164.45 175.20 183.62 194.70 206.06
213.87 219.55
SPACE 3
MOVE 1,5 18X1 TO 1,1 (INDEX FIVE)
NOTE$WORKSHEET
MPRINT 1,1 18X6
MPRINT 1,17 18X2
MPRINT 1,30 18X1
NOTE1 $ RHO, COEFF, DURBIN WATSON
DEFINE 1.0 IN COL 13
RESET 18
BEGIN
```

Initialize Rho, ρ

```
1ADEFINE 1,8 18X1 TO B E 0.80 ( RHO)
2SQUARE 8 PUT IN 9
3SUBTRACT 9 FM 28 PUT IN 10
4SQRT OF 10 IN 11
```

Transform composite Price Index

```
5AMULT 1,1 18X1 BY 11 PUT IN 1,7
6AMULT 1,1 17X1 BY 8 PUT IN 2,7
7SUBTRACT 7 FM 1 PUT IN 20
```

Transform GNP_d

```
8AMULT 1,30 18X1 BY 11 PUT IN 1,31
9AMULT 1,30 17X1 BY 8 PUT IN 2,31
10SUBTRACT 31 FM 30 PUT IN 21
```

Transform constant

```
11AMULT 1,13 18X1 BY 11 PUT IN 1,12
12AMULT 1,13 17X1 BY 8 PUT IN 2,12
13SUBTRACT 12 FM 13 PUT IN 22
```

Set up Row One

```
14MOVE 1,7 1X1 TO 1,20
15MOVE 1,31 1X1 TO 1,21
16MOVE 1,12 1X1 TO 1,22
```

Regress

```
17SFIT Y IN 20 WTS 1.0 2 VR IN 22,21 C IN 33 R IN 34
```

Calculate Durbin Watson Statistic

```
18MOVE 2,34 17X1 TO 1,35
19SUBTRACT 34 FM 35 PUT IN 35
20SQUARE 34 PUT IN 36
21SQUARE 35 PUT IN 37
22SUM 36 ROW 1 TO 18 PUT IN 38
23SUM 37 ROW 1 TO 17 PUT IN 39
24DIVIDE *1,39* BY *1,38* PUT IN 40
```

Print: Rho, Coefficients including S^2, Durbin Watson Statistic

```
25SPACE 3
26PRINT NOTE
27SPACE 3
28MPRINT 1,8 1X1
29SPACE 3
30MPRINT 1,33 10X1
31SPACE 3
32MPRINT 1,40 1X1 (DW)
```

Increment Rho

```
33INCREMENT INSTRUCT 1 BY 0,0,0,0,0.01
  FINISH
  PERFORM 1,33 19 TIMES
```

Next Composite Index

```
RESTORE 1 TO 1,8,18 ,1,0.80
MOVE 1,6 18X1 TO 1,1
PERFORM 1,33 19 TIMES
STOP
```

Vita

Thomas P. Lennertz was born on June 26, 1947 in Chicago, Illinois. He graduated from Andrean High School in Merrillville, Indiana in 1965. He attended the University of Notre Dame. During the summer months he held a series of jobs ranging from steel mill laborer to research engineering technician. He graduated from Notre Dame in 1969 with a Bachelor of Science degree in Aerospace Engineering. At the same time he received his commission in the United States Air Force through the Reserve Officers Training Corps. After 5-1/2 months of menial jobs in the steel mill, he began active duty in the Air Force. He served 5-1/2 years in the Air Force all as a Space Object Identification Analyst in Florida, Alaska, and Colorado.

Permanent Address: 530 West 56th Place
Merrillville, Indiana 46410

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER (14) GOR/SM/76D-9	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER (9) Master's thesis,
4. TITLE (and Subtitle) (6) AN INVESTIGATION OF THE FORECASTING OF AERONAUTICAL PRICE INDICES.		5. TYPE OF REPORT & PERIOD COVERED MS Thesis
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) Thomas P./Lennertz ~~Captain, USAF~~		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Air Force Institute of Technology (AFIT-EN) Wright-Patterson AFB, Ohio 45433		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS AFIT-EN		12. REPORT DATE (11) Dec[redacted]76
		13. NUMBER OF PAGES 112 (12) 115p.
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribtuion unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Approved for public release;
IAW AFR 190-17

Jerral F. Guess, Capt. USAF
Director of Information

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Price Indices
Cost Indices
Escalation
Autocorrelation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

An investigation of the forecasts of aeronautical price indices done by the Aeronautical Systems Division (ASD) of the United States Air Force Systems Command was performed. The state of the art in correction for autocorrelation was summarized. The historical price indices used by ASD were recalculated. Using these the recalculated ~~historical~~ composite price indices, different forecasting techniques were examined, including: simple linear re-

(Cont on p 1473B)

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified 012225

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(cont fr P 1473A)

gression, multivariate regression, regression with nonlinear variables, Generalized Least Squares with the first order autoregressive model, first differences, and per cent change regression. The Gross National Product price deflator, the Primary Metals Index, the Electrical Machinery Index, the national total unemployment rate, and the money supply, M1, were all tried as independent variables in the regressions. The Durbin Two Step, the Search technique, and a Bayesian technique were all used to estimate the first order autoregressive coefficient, ρ. When starting with the composite price indices, the per cent change regression, with the Gross National Product price deflator as the only independent variable, was recommended as the best forecasting technique. Prediction intervals were calculated. The effect of errors in the Wharton Econometric Forecasting Associates prediction of the Gross National Product price deflator was discussed.

1473B